



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 162983

TO: Phuong Bui  
Location: REM-2A15&2C18  
Art Unit: 1638  
Monday, August 29, 2005

Case Serial Number: 10/677179

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Bui,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

BEST AVAILABLE COPY

**THIS PAGE BLANK (USPTO)**

	Query Match	100.0%;	Score 1626;	DB 6;	Length 1626;
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Gaps 0;
	Matches 1626;	Conservative	0;	Indels 0;	Indels 0;
Qy	1	GCACGAGCTCAGTTCGTCTCTCGCCGGCGGCGAGCAACTATCATCTTCATTAG	60		
Db	1	GCACGAGCTCAGTTCGTCTCTCGCCGGCGGCGAGCAACTATCATCTTCATTAG	60		
Qy	61	CTCATCCAATCTATTCCGATGACGCCGTGCCAAGCCACAGCAGTGCGCGCACTCACGG	120		
Db	61	CTCATCCAATCTATTCCGATGACGCCGTGCCAAGCCACAGCAGTGCGCGCACTCACGG	120		
Qy	121	GCACGGCTGCACCCC CGCGGATTCGCGCGCTTGCGAGTTTGGCCCCAGCCTCTCCTCC	180		
Db	121	GCACGGCTGCACCCC CGCGGATTCGCGCGCTTGCGAGTTTGGCCCCAGCCTCTCCTCC	180		
Qy	181	CTCCGCTTCGCGTGCACCGCTGCCGCACTGCTCGCTAGAGGTGAAGCATCTGGAAC	240		
Db	181	CTCCGCTTCGCGTGCACCGCTGCCGCACTGCTCGCTAGAGGTGAAGCATCTGGAAC	240		
Qy	241	ACGTTTTGGAACACTTTTCAGGTTGCAACTATGTTGAATCTCATGGGGGTGGTGTGGT	300		











Db	538	GGAGTGAGATCTGTACAGGGGGTGTAGATCATCAGCAAGAGAGACTATTGGGAGAGTT	597
Qy	634	GCTGAGGGGCCCTCCCAAGAAATCTTAAGCTCAAAATGTGATGATAGAGATCTTTGTCG	693
Db	598	GCTGCTGAGCAGTGTGCTAAGAAATCTCAAACTTATTCTGGAGCTGAGGTTCTTGCT	657
Qy	694	TTTGTTCCTCAAGTGCATCAGTGTGCTCCAGAGACGGGTTGATTATGGGCTCTGTA	753
Db	658	TATGTTTCTCAAGTTCACCAAGTTGTACTTCTGAGGATTTGATTGATCATCAGAACGTG	717
Qy	754	ACTTTGGAAACAGATAGAGACCAACATCGTTAGATGTCTCTGATCCAGAGTACGCGAGAGAG	813
Db	718	ACTTTAGAGCAGATAGAAAGCAATATTGTCGAGTCCCGGATCCGGATATGCGAGAGAG	777
Qy	814	ATGATAGAGCAATGCACAGAGTACGAGTTGAGGGGATTCGGTCGGTGGAGTATCAACA	873
Db	778	ATGATTGCTGCTATTGATGCTGTACGAGTGAGAGAGACTCTGTTGGTGGTGTGTA	837
Qy	874	TGCGTCGCTAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTGTTTCGACAACTCGAATCC	933
Db	838	TGCATTGTTAGAAATCTCCACGGGGTCTTGTTACACCACTCTTCGATAAACTTTGAAGCT	897
Qy	934	GAACTGGCAAAAGCTATGCTTTCTATTCTCGAGCAACGGGTTTCGAGATTGGCAGCGGA	993
Db	898	GAGCTAGCCAAAGCTTGCATGTCTATTCTCGCAAAAGGGTTTGAGTTTGGCAGTGGC	957
Qy	994	TTCCGGCGGACCGACTTCGACAGGAAGTGAAGATATGATGATGTTTATATGATGAAGCT	1053
Db	958	TTTTCAGGACCGTTTCATGCTGTGACGAGCATATGATGAAATCTATATGATGATGAGCAT	1017
Qy	1054	GGAGTGTGACGACGCACTAACTGCTCGGGTGTGTCGAGGAGGGGATATCGAATGTT	1113
Db	1018	GCCGAATCAGCAAGAACCAACAGATCTGTGTGAATCCAGGGCGGTATATCAAAATGGA	1077
Qy	1114	CAGATTGTGCATCTCAAAAGTTGCTTTTAAGCGCACCACTATCTCGGGGTGAAACAGAAC	1173
Db	1078	GAACTTATCAATATGAGATAGTTTTCAGCCAACTTCAACTATTCTTAGGAAGCAGCA	1137
Qy	1174	ACCGTGTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAGGGCGCCATGACCCATGC	1233
Db	1138	ACTGTGACAGAGACAAAACAGAAACAGAACTCATCGCTAGGGGTGCGCATGCTTTGT	1197
Qy	1234	GTCCGCCCTCGAGCTGTTCTGTGTGGAATCCATGGCGGTTGGTCTCATGAGACAG	1293
Db	1198	GTGGTTCGCCAGCTGTTCCAAATGTTGAAGCAATGGTCGCTTGGTCTGTTGATCAG	1257
Qy	1294	CTGATGGCGCAGTGGCTCAGTGCAGAGATGTTGGCGCTCAATACTGCACTTCAAGAACC	1352
Db	1258	TTTATGGCTCAGATTTCACAGTGATGATGTTCCCAATCAATCTCTGNACTACAGNAAC	1316

RESULT 6  
LECHOSYNB  
LOCUS LECHOSYNB 1802 bp mRNA linear PLN 06-FEB-1996  
DEFINITION L.esculentum chorismate synthase 2 precursor.  
ACCESSION Z21791  
VERSION Z21791.1 GI:410483  
KEYWORDS chorismate synthase 2.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE  
1 (bases 1 to 1802)  
Gorlach,J., Schmid,J. and Amrhein,N.  
Differential expression of tomato (Lycopersicon esculentum L.)  
genes encoding shikimate pathway isoenzymes. II. Chorismate  
synthase  
JOURNAL Plant Mol. Biol. 23 (4), 707-716 (1993)  
MEDLINE 94072719  
PUBMED 8251624  
REFERENCE 2 (bases 1 to 1802)

AUTHORS Gorlach,J., Raesecke,H.R., Abel,G., Wehrli,R., Amrhein,N. and Schmid,J.  
TITLE Organ-specific differences in the ratio of alternatively spliced chorismate synthase (LeCS2) transcripts in tomato  
JOURNAL Plant J. 8 (3), 451-456 (1995)  
MEDLINE 96045551  
PUBMED 7550381  
REFERENCE 3 (bases 1 to 1802)  
Gorlach,J.  
Direct Submission  
Submitted (22-FEB-1993) Joern Goerlach, Institute of Plant Sciences, Swiss Federal, Institute of Technology, Sonnegstr. 5, Zurich, CH-8092, Switzerland  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..1802  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/strain="UC828"  
/db\_xref="taxon:4081"  
/clone="cLeCS2"  
42..1337  
/EC\_number="4.6.1.4"  
/codon\_start=1  
/product="chorismate synthase 2"  
/protein\_id="CAA79854.1"  
/db\_xref="GI:410484"  
/db\_xref="COA:Q42885"  
/db\_xref="UniProt/Swiss-Prot:Q42885"  
/translation="MASSMLTKQFLGAPFSGGQOPSKLSSNLRFPHRSOPKRL  
EIQAGNTFNGYFRVTFGSHGGVCIIDGCPRLPLSESDMQVELDRRFGQSR  
TTPRKTDTCKISSGTADGLTGTSPIKVEVPNTDQGNDSYSEMSLAYRPSHADATYDF  
KYGVSQGGSSARETIGRVAAGAVAKILKLYSGTEILAVSVQVHVLPEDLVD  
NQIUTLEQIESNIVRCNPPEYAEKMGADYVVRGDSVGGVVTCTIVRVPRGLGTPV  
FDKLEAELKACMSLPATKGFEGFAGFTGSEHDEFFMDEHDQIINTKRSGG  
IQGGINGEIINNRVAFKPTSTIARKQHTVSRDKHETELIARGHDPQVVPRAVPMVE  
AMVALVLDQLMTQYACMLFPVNLTLQEPLOPSTTKSA"  
42..185  
sig\_peptide  
186..1334  
mat\_peptide  
/product="chorismate synthase 2"  
/EC\_number="4.6.1.4"  
1802  
polyA\_site  
ORIGIN  
Query Match 35.3%; Score 574.6; DB 8; Length 1802;  
Best Local Similarity 67.9%; Pred. No. 1.5e-152;  
Matches 802; Conservative 0; Mismatches 379; Indels 0; Gaps 0;  
Qy 172 TCCTCTCTCCCTCGCGTTGCGCTGCAACCGCTGCGCGCACTGCTCGCCTAGAGGTGAAGCA 231  
Db 126 TCTTCAAACTCCGATTCCCTACTCACCGATCAAGCCCAACGACTAGAGATACAAGCT 185  
Qy 232 TCTGGAACAGCTTTGGAACACTACTTTCAGTTGCAACCTATGGTGAATCTCATGGGGGT 291  
Db 186 GCTGGTAATACATTGGAATAATTATTTCCGCGTTTACACGTTTGGAGAATCTCAGCGTGT 245  
Qy 292 GGTGTTGCTGTGTATTCAGTGGTTGTCACCTAGAAATTCACCTCAGTGGAGCAGACTA 351  
Db 246 GGTGTGCGTTGTATATTTGATGATGTGTCCTCGACTCCACTCTCAGAGTGGATATG 305  
Qy 352 CAAGTTGAATTCGATCGAAGACGGCCCGGACAGAGCAAGATAACCTCCACAAGGAAGAG 411  
Db 306 CAAGTGAACCTTGACAGGAGGAGCCAGGTCAAAGCCGAATTACCACACCAAGAAAAAG 365  
Qy 412 ACTGATACATGCMAAATTCGTGAGGACACATGAAGGGGTGACTACTTGGAAACCCAAATT 471  
Db 366 ACTGATCTTGCAAAATATCATCAGGCACCTGCTGATGGTTTGACTACGGGATCTTCCAATC 425  
Qy 472 CTGTTATTGTCCTCCAAACACAGATCAATAGGCGATGATCACCGTGAATAGGCAATGTG 531  
Db 426 AAGTTGAAGTACCAACACTGACACAGAGGAATGACTATAGTGAATGTCCCTTGCT 485  
Qy 532 TACCGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACCGGTGTTAGAGCTGTACAG 591



```
Db      804  GTCTCTCAAGCCCAAGAGTTGTTCTTCCAGAGGGTTTGGTTGATCATGAGACTATCA 863
Qy      757  TTGGAACAGATAGAGAGCAACATGTTAGATGCTCTGATCCAGAGTACGACAGAAAGATG 816
Db      864  CTGGAACAGATTGAAGTAACATCTTAGTGTGCCAGATTGAGATATGAGGAGAGATG 923
Qy      817  ATAGAGCAATCGACAGAGTACGAGTTGAGGGGATTCGATCGGTGGAGTATCATATGC 876
Db      924  ATTGCTGCTATTGATGCTGTCCGAGTGAAGGAGACTCTGTGTGGAGTTGTGACATGT 983
Qy      877  GTCCGTAGAAAGCTTCTCGCGGGCTCGGTCTCTCGTGTTCGACAAAGCTCGAATCCGAA 936
Db      984  ATTATGAGGAGAGTCCAGGTGGGCTTGGTTCACCAAGTATTGATTAACCTTGAGGCTGAG 1043
Qy      937  CTGGCAAAAGCTATGCTTTCTTATCTCGAGCAACCGGTTGAGATTTGGCAGGGATTC 996
Db      1044  CTTGCCAAAGCGTGATGTGCTATTACAGCAACAAAGGGGTTTGAATTTGGGAGTGAATT 1103
Qy      997  GCGGGACCGACTTTCACAGGAGTACGATTAATGATGATGATTTATATGATAGGCTGGA 1056
Db      1104  TCAGGCACATTTCTGACTGGTAGTGAACATATGATGATGATTTCTACACAGACGAGATGGA 1163
Qy      1057  AGTGTACAGACACGACTAATCGCTCGGTGTGTGCGGAGGAGATATCGAATGTTGAG 1116
Db      1164  AGAATCAGGACTAGAACAAATCGATCTGGTGGATACAGGGTGGTATATCCACGGAGAA 1223
Qy      1117  ATTTGCACTTCAAAAGTTGCTTTTAAAGCCGACACCATCTATGGGGTGAACACAGAACCC 1176
Db      1224  ATCAATAAATGAGATAGCTTTCAAGCCGACATCTACAATTTGGAAGAACGAGAACTACT 1283
Qy      1177  GTGTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAGGCGCCATGACCCATGGCTC 1236
Db      1284  GTGACAGAGAACGTTGAAGAGATTTGAATCTATAGCACGTGGTCGCCACGACCTTTGTGT 1343
Qy      1237  GCGCTTCAGAGCTGTTCTGTGTGGTGAATCCATGCGCGGTGGTCTCATGACCGAGCTG 1296
Db      1344  GTACTTCGAGCTGTTCCAAATGGTGGAGCTATGGTAGCCCTGGTCTCTGATCACTG 1403
Qy      1297  ATGGCGACGTTGGCTCAGTGGAGATTTGCGGCTCAATATGCACTTCAAGAACCC 1352
Db      1404  ATGCTTCAACATGCTCAGGGCAACCTTTTCTCATCAATCTGCTTCAAGAACCC 1459
```

```
RESULT 8
LOCUS      CQ805346                1311 bp      DNA      linear      PAT 10-MAY-2004
DEFINITION Sequence 1757 from Patent WO2004035798.
ACCESSION  CQ805346
VERSION    CQ805346.1  GI:47111307
KEYWORDS
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE  1
AUTHORS   Inze, D., de Veylder, L. and Vlieghe, K.
TITLE     Identification of novel e2f target genes and use thereof
JOURNAL   Patent: WO 2004035798-A 1757 29-APR-2004;
CropDesign N.V. (BE)
FEATURES   Location/Qualifiers
            1..1311
            /organism="Arabidopsis thaliana"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3702"
```

```
ORIGIN
Query Match      32.9%; Score 535.4; DB 6; Length 1311;
Best Local Similarity 66.9%; Pred. No. 2.3e-141;
Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy      216  CCTAGAGGTGAAGGCATCTGGAAACACGTTTGGAAACTACTTTTCAGGTTGCAACCTATGG 275
```

```
Db      135  CTTCCAGATCAAGCTACTTGGAGTTTCAATGGGACTCATTTTCGAGTTTCAACTTTGG 194
Qy      276  TGAATCTCATGGGGTGGTGTGGTTGTGTATCATGATGGTGTGCCACCTAGAAATTCACCT 335
Db      195  AGAATCATATGGAGGAGGTTGGTTGTATCAATGATGGTTGTCTCTCTGATTCACCT 254
Qy      336  CACTGAGGACAGCTTACAGTTTGAATCTCGATCGAAGACGCCCGGACAGAGCAGAAATAC 395
Db      255  TACTGAATCTGAATTCGAATTCGATCTCGATAGAGAGCGCTGGTTCAGAGCAGGATCAC 314
Qy      396  CTCCACAAGCAAGAGAGACTGATACATGCAAAATTTCTGTCAAGGACACATCAAGGGGTGAC 455
Db      315  AACTCTTAGAAGAGAGACTGATTAATTCGCCGATATCGTCTGGAGTCTCTGAGGAATGAC 374
Qy      456  TACTGGAAACCCCAATCTTGTATTTGTCCAAACACAGATCAAAATAGGAGTATCATCCG 515
Db      375  GACAGGAACACCTATCCATGTGTTGTACCAACACAGATCAGAGAGGACTTGAATTACAG 434
Qy      516  TGAATAGCCAAATGTGTACCGACTTCTCATGACAGACGCAACTTATGACTTCAAGTACGG 575
Db      435  TGAATGTCCGTTGCCCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAGTATGG 494
Qy      576  TGTTAGAGCTGTACAGGGAGGTGGGAGGTCTCGGGCAGAAACCCGTTGGAAGGGGTGGC 635
Db      495  TGTCAGATCAGTGCAGGGTGGAGGAGATCTTCAGCTAGAGAGACCAATTTGGAGAGTTGC 554
Qy      636  TGCAAGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGGGATTTAGAGATCTTTGCTT 695
Db      555  TCCTGGAGCTTTGGCCAAAGAAAAATTTTGAAGCAATTTGACAGGAACCTGAGATCTTCGCTA 614
Qy      696  TGTTCCAAAGTGCATCAGGTTGTCTCCAGAGACGCGGTTGATTTATGGTCTGTAAAC 755
Db      615  TGTCTCCAAAGTTCCACATGTGTACTTCCAGAAAGAAATTTGTAGACCAACAGAAATTTAAC 674
Qy      756  TTTGGAACAGATAGAGAGCAACATCGTTAGATGCTCTGATCCAGAGTACGACAGAAAGAT 815
Db      675  ACTCGAACAGATAGAAATTAACATTTGTCAGATGCCCTAATCCCGAGTATGGGAAAGAT 734
Qy      816  GATAGACGAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTATGATCATG 875
Db      735  GATAGCTCGCAATGATGCTCTGACAGCAAAAGGGAACTCTGTTGGTGGTGTGTGTGACCTG 794
Qy      876  CGTGCCTAGAAAGCTTCTCGCGGCTCGGTTCTCTGTTCTCGACAAAGCTCGAATCCGA 935
Db      795  CATTTGTTGGAAAGCTCTCAAGTGGGCTTGGTACACCGGTTTTCGATAAATCTTGAGCAGA 854
Qy      936  ACTGGCAAAAGCTATGCTTTCTATTCTCGAGCAACCGGTTTCGAGATTTGCGACGGATTT 995
Db      855  ACTGGCAAAAGCTTGTATGTGCTACCTGCAACTAGGGATTTGAGTTTGGNAGCGGCTT 914
Qy      996  CGCCGGACCGACTTGAACAGAGTGAACATATGATGAGTTTATATGATTAAGGCTGG 1055
Db      915  TGCAGGTACCTTTTGTGCTGTCTTGAAACAAATGATGAGTTCTATACCGATGAAATGG 974
Qy      1056  AAGTGTACAGACAGGACTAATCGCTCGGTGGTGTGACGGGAGGGATATCGATATTTGA 1115
Db      975  AAGAATACGTAACCAAGAACCAACCGATCTGTGGAATTTACGGGAGGGATCTCAAAATGTGA 1034
Qy      1116  GATTGTGCACTTCAAAAGTGTCTTTTAAGCCGACACCATCTATCGGGGTGAAACAGAAACAC 1175
Db      1035  AATTAATAAATCATGAGAGTAGCTTCAAGCCCAACATCAACAAATTTGAAGGAAGCAAAATAC 1094
Qy      1176  CGTGTCAAGGAGCGGTGAGAACGTTGAGCTTTAGCAAGAGGGGCCCATGACCCATCGGT 1235
Db      1095  GGTAAACAGAGACAAAGGTAGAAACCGAAATGATTTGCGGTGTGTCTCATGATCTCTGTGT 1154
Qy      1236  CGCCCTTCGAGCTGTTCTCTGTGTGGATCCATGGCCCGGTTGGTCTCATTGGACACGCT 1295
Db      1155  TGTTCACAGAGCTGTGCAATTTGGTGAAGCAATGTGTGCTCTAGTTCTTTGTGGATCAATT 1214
Qy      1296  GATGGCGCACTGTGCTCAGTGCAGAGATGTTTCGCGCTCAATACTGCACTTCAAGAACCC 1352
```





Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Arabidopsis cdna clones  
Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 1603)

Chen,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE  
JOURNAL

Direct Submission  
Submitted (18-SEP-2001) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs (Chen,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Chen,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1..1603  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="RAFL09-47-O14 (R14512)"  
/ecotype="Columbia"

5'UTR

CDS

1..122  
123..1433  
/note="putative chorismate synthase"  
/codon\_start=1  
/product="At1g48850/T24P22\_3"  
/protein\_id="RA109759.1"  
/db\_xref="GI:1598284"  
/translation="MASSSLTSKSLGSLKSSLSPELRRLSSPAVQISLRQTRK

NFIQATGSSYTHFRVTFGESHGGVGIIDGPPRIPLTESDLQDLDLRRPQGS  
RITPRKEDTCRISSEGEVGTPIHVFPNTDQRLDYSMSVAYRSHADATY  
DMKVGVSQGGSSRSARETIGRVAPGALAKKILKQFAGTEILAYVSQVHVVLPEEL  
VDHENLTLEQENNTVRCNPPEYAEKMIADAVRTKGNSSVGVVTCIVRNAPRLGT  
PVFDKLEALACWMLPATKGFPEFGSGFAGTFLTGLEHNDDEFYTDENGRIRTRTNS  
GVIQGISNGEILINRVAFKPTSTIGRKQNTVTRDKVETEMIAGRHDPVVRPAPVM  
VEAMVALVLVDQLMAQYAQCHLFFINPELQBPQLQIEQPQNATAL"

3'UTR

ORIGIN

Query Match 32.9%; Score 535.4; DB 8; Length 1603;

Best Local Similarity 66.9%; Pred. No. 2.4e-141;

Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 216 CCTAGAGGTGAAGGCATCTGGAAACACACGTTTGGAAACTACTTTACAGTTGCAACCTATGG 275

Db 257 CTTCCAGATACAAAGCTACTGGAAGTTTCATATGGGACTCATATTTTCGAGTTCCTCAACTTTGG 316  
QY 276 TGAATCTCATGGGGTGGTGTGGTGTGTATCAGTGGTTGTCCACCTAGAATTCACACT 335  
Db 317 AGAATCATCATGGAGGAGGTGGTGTGTATCATTTGATGGTTGTCTCTCCTCGTATTCACACT 376  
QY 336 CACTGAGGCAGACCTACAAGTTGAACCTCGATCGAAGACGCCCGGACAGACAGAGCAGAATAAC 395  
Db 377 TACTGAATCTGATTTGCATTCGATCTCGATAGAGGAGGCCCTGGTCAGACGAGGATCAC 436  
QY 396 CTCCACAAGGAAGAGAGACTGATACATGCAAAAATTTCTGTCAAGGACACATGAAGGGGTGAC 455  
Db 437 AACTCTTAGAAAAGAGACTGATACTTGGCCGATATCGTCTGGAGTCTCTCTGAAGGAATGAC 496  
QY 456 TACTGGAAACCCAAATTTCTGTGTATTTGTTCCAAAACACAGATCAATATAGGAGTGTATCACCG 515  
Db 497 GACAGGAACACCTATCCATGTGTGTATGTTGTAACCAACACAGATCAGAGAGGACTTGTATTACAG 556  
QY 516 TGAATAGCCAATGTGTACCGACCTTCTCATGACGACGCAACTTATGACTTTCAGGTACGG 575  
Db 557 TGAATGTCCGTTGCCCTATAGACCATCGCATGCTGATGCAACTTATGATGAAGATATGG 616  
QY 576 TGTTAGAGCTGTACAGGGAGGTGGAGGTCTCTCGGCAGAAAAACCGTTGGAAGGGTGGC 635  
Db 617 TGTCAAGTCAAGTGCAGGGTGGAGGAGATCTTCAAGCTAGAGACCATTCGAGAGTTGC 676  
QY 636 TGCAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTT 695  
Db 677 TCTTGGAGCTTGGCCCAAGAAAAATTTTGAAGCAATTTTGCAGGAACTGAGATTTCTTGCTA 736  
QY 696 TGTTTCCAAAGTGCATCAGGTGTGCTCCAGAAAGACGCGTTGATTTATGCTGGTCTGTAAC 755  
Db 737 TGTCTCGAAGTTCCACATGTTGTACTTCCAGAAAGATTTGTTAGACCAAGATTTTAC 796  
QY 756 TTTGGAAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGACAGAGAAT 815  
Db 797 ACTGAAACAGATAGAAAAATTAACATTTGTCAGATGCCCTAATCCCGAGTATCGGAAAAAGAT 856  
QY 816 GATAGAGCAATCGACAGATACAGTTTCAGGGGATTCGGTGGTGGAGTATCATCATG 875  
Db 857 GATAGCTGCAATTCATGCTGTCTAGGACAAAAGGAACTCTGTTGGTGGTGTGTGACCTG 916  
QY 876 CGTCTGTAGAAAACTTCTCCCGGGCTCGGTCTCTCTGTGTTCGACAAGCTCGAATCCGA 935  
Db 917 CATTTGTCGAATGCTCCAGTGGGCTTGGTACACCGTTTTTCATTAATTTGAGAGAGA 976  
QY 936 ACTGGCAAAAGCTATGCTTTTCTATCTTCGACGACAAACGGGTTTCGAGATTCGACGCGATT 995  
Db 977 ACTGGCAAAAGCTTGTATGTCTGCTACCTGCAACTAAGGGATTTGAGTTTGGAAAGCGCTT 1036  
QY 996 CGCCGGGACCGACTTGAACAGAAAGTGAACATATGATGAGTTTATATGATAGGCTGG 1055  
Db 1037 TGCAGGTACCTTTTGTGACTGCTTCTTGAACCAATGATGAGTTCTATACCGATGAAATGG 1096  
QY 1056 AAGTGTCAAGACCGGACTAATCGCTCGGTGGTGTTCGAGGAGGATATCGAATGTGA 1115  
Db 1097 AAGNATAGTACCAAGAACCAACCGATCTGTTGAAATTCAGGAGGAGTCTCAATGTGTA 1156  
QY 1116 GATTGTGCACTTCAAAGTGTCTTTTAAAGCCGACACCATCTATCGGGGTGAAACAGAACAC 1175  
Db 1157 AATAATAAACATGAGAGTAGCTTCAAGCCAAACATCAAAATTTGGAAGGAAGCAAAATAC 1216  
QY 1176 CGTGTCAAGGAGGCTCAGAACGTTTGAGCTTCTAGCAAGAGGGGCCCATGCCCTCGCT 1235  
Db 1217 GGTAAACAGAGACAAGGTAGAAACCCGAAATGATTCGCGGTGGTGTGTCATGATCTCTGTGT 1276  
QY 1236 CGCCCTCTGAGCTGTTCTCTGTGTGGAATCCATGGCCCGCTTGGTCTCTATGGAACAGCT 1295  
Db 1277 TGTTCACAGAGCTGTGCCAATGGTGAAGCAATGGTGGCTCTAGTCTTCTTGGATCAATT 1336  
QY 1296 GATGGCCACGCTGGCTCAGTGCAGAGATGTTTCGGCTCAATCTGCACTGCACTTCAAGAAC 1352  
Db 1337 GATGGCCCAATAGCAACATGCCATTTGTTTCCAAATAATCCAGAGTTGCAGGAAC 1393



QY	Db	1351	CGATGGGCTC	1361
Db	725	CCAATCGGCTC	735	
RESULT 12				
BT009607				
LOCUS				
DEFINITION				
BT009607				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
ORIGIN				
Query Match				
Best Local Similarity				
Matches				
Conservative				
Indels				
Gaps				
0;				
631	GTGGCTGCAGGGGCCCTCCCAAGAAAATCTTAAAGCTCAAAATGTGGATTAGAGATCTTG	690		
Db	5	GAGGCTGCAGGAGCTGTTCGAAAGAAAATCTTAAAGCTGAAATGTGGATTAGATCTTA	64	
QY	691	TCGTTTGTTCGAAAGTGCATCAGGTTGTCTCCCAAGAACGCGGGTTGATTATGGGTCT	750	
Db	65	GCATTTGTTCGAAAGTGCATCAGGTTGTCTCCCAAGAACGCGGGTTGATTATGAAACT	124	
QY	751	GTAACCTTTGGAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGCAGAG	810	
Db	125	CTTACCTTGCATAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGCAGAG	184	
QY	811	AAATGATAGACGCAATCGACAGAGTACGAGTTCCAGGGGATTCGGTCCGTTGAGTATGATC	870	
Db	185	AAATGATGATGCAATTCGAAAGTACGAGTTTAAATGGGAATTCGATTTGGTGGGTGGTC	244	
QY	871	ACATGGCTGCTAGAAAGCTTCTCGGGGCTCGGTTCTCTGTTCTCGACAAAGTTCGAA	930	
Db	245	ACATGCAATTCGCCAGAAATGTCTCGTGGGCTTGGCTCTCTCTGATTTTGAACAACTGAA	304	
QY	931	TCGGAATCGCAAAAGCTATGCTTCTATTCCTCGGAGCAACGAGTTTCGAGATGGCAGC	990	
Db	305	GCTCTACTGGCAAGGCTATGCTTCTCTCTCGCAAGCAAGGGGTTTGAGATCGGTAGT	364	
QY	991	GGATTCGCGGGACCGATTCGACAGAGTACGAGTTCGAGGGGATTCGGTTCGAGTATG	1050	
Db	365	GGATTTGCAAGTACTGACCTTAAGTGGAGTACGAGTACGAGTTCGATTTATGACGAG	424	
QY	1051	GCTGGAAGTGTACGAGCAACGAGTTCCTCGGGGCTCGGTTCTCTCGTGTTCGACAAAGCTCGAA	1110	
Db	425	GCTGGAATGTAAAGAACACGAGCAACCAATCGCTCGGGGCTTCGGCTCTCTCTGATTTGACAAACTTGAA	484	
QY	1111	GTTGAGATTTGCACTTCAAGTTCGTTTAAAGCCGACCAATCTATCGGGGTGAACAG	1170	
Db	485	GGTGAACCTATATACCTTCAAGTGTAGCTTTCAAGCCACAGCAACTATTTGGGGAAGACAA	544	
QY	1171	AAACACGTTGCAAGGAGGCTCAGAACGTTGAGCTTCTAGCAAGAGGCGCCATGACCCA	1230	
Db	545	AATACCTGTAACAGGGATCATGAGGATTCGAACTTCTGCAAGGGGTCGCATGACCCA	604	
QY	1231	TGGCTGCCCTCGAGCTGTCTCTGTGTGGTGAATCCATGGCGCGTCTTGGTCTCTCATGGAC	1290	
Db	605	TGTGTGCTCCCTCGGGCTGTTCCAATGGTGGAGAGCATGGCTGCATTTGGTCTCTCATGGAC	664	
QY	1291	CAGCTGATGGCGCAGCTGCTCAGTGCAGAGATGTTTCGGGCTCAATATGCACTTCAAGAA	1350	
Db	665	CAGCTGATGGCGCAGCTGCTCAGTGCAGAGATGTTTCGGGCTCAACCTCGGCTCAAGAA	724	



```
source
1. .339650
/organism="Nostoc sp. PCC 7120"
/mol_type="genomic DNA"
/db_xref="taxon:103690"
/notes="synonym:Anabaena sp. PCC7120"
complement(98. .2698)
/gene="all0594"
complement(98. .2698)
/gene="all0594"
/notes="ORF_ID:all0594"
/codon_start=1
/transl_table=11
/product="aminopeptidase"
/protein_id="BAB72552.1"
/db_xref="GI:17129940"
/translation="MPHSYSFDODSNAHKSPELPGARPHYNDRPGQVHEIFLDNLNLD
IQQSYEGSCSRLPIRNGIDRLTLDAVNLQIRSVQVDEVGQDFDYDGEQVLIQLSQ
PTEIGRIIDIAAYAEKPGQRIYFIQPSDHYPNKTVQVMTQGBEDSRFPWPCFDYP
GQLSTSEIRVRYPKSLMAISNGELIDTQBDGKHKIYHMSQQQIHPTIYLTMLAVADFAE
IRDEWRGKPTVYVSKGREADAKRSMGKTPRMIEFLSEKYGAYAPPKYAOVCVSDFI
FGCMGNTSTLLTDRCLDRAALDNRNTESLVHLELAHQWFGDILLVHKHSHAWIKE
GNASYEVMWTEQYESAEAAVYRLLEARSYLSSESGRYRRPMWTHVREALIELYDRH
TYEKSCVYHMRABHGEDLFWQAQTFVRDNAHKIVETIDLRRAIEKATGRNLAFLF
DQYVRGHPDPKAYSWDGEAKLAKVTVTQVQPADKQVDFDLRIPIGFQYNQTEA
AAKQTFVVRHERBQSIFYPLETEKPDFISFDVGNHMLKTVLSLEYPIPELKAQLASDP
DPSIRYAAALAKKGLEATKALAAKNEBFWGVVRVEVAKGLAEIKLDOAFDGLVE
GLADNAYVRRVAQALAKIKTLASVAKVGVKQGDASYVVEASACLAGAIAINF
DPDKSEKAKIKLQSVLERAGNEVRSGLAGGLAELKTSALDLILEYTKLGVPQ
PURLAADQADGKISQGSVPNULRIIARIGELSKENFFLTLMAVVAALGQMETPKAMG
VURSLADQADGKISQGSVPNULRIIARIGELSKENFFLTLMAVVAALGQMETPKAMG
LEAKSQYVKQ"
complement(2862. .3059)
/gene="aol0595"
complement(2862. .3059)
/notes="ORF_ID:aol0595
similar to cytochrome d oxidase chain I"
/codon_start=1
/transl_table=11
/protein_id="BAB72553.1"
/db_xref="GI:17129941"
/translation="MRHCVHPILLSLPLYLKIIILNIQWKPSDRPHMVGLIYAPRVN
IAIGPFLAGMLVSTLFEVR"
complement(3151. .3810)
/gene="all0596"
complement(3151. .3810)
/notes="ORF_ID:aol0596
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72554.1"
/db_xref="GI:17129942"
/translation="MGVCHSKDVSBFCLEGRPFMEFVYKDGVKLKGILLLLTADGECYVK
LSKQLRFTFWRPLTQLQVGEKKYDAKTGVTLKABAVMAARTETSSAKQVPMAP
ENTKPKPEKAKATILVCOKSCDMKRGKALCOALEATISDRGLEQDQVTKGTGCKWNC
KAGPNLVMPDKTRYTRIQAQVYPLMKHMFADSQQSDSHSQAVISNTNKFEPFTRVSA
F"
complement(3877. .4134)
/gene="aol0597"
complement(3877. .4134)
/gene="aol0597"
/notes="ORF_ID:aol0597
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72555.1"
/db_xref="GI:17129943"
/translation="NKKQSSVEEKSVEVVDWRWQVYQRLQELLEIPWCNEANQPLK
VELVTPMTVQWLSVMRQFTSSRQDLIYSLEINWOSRYPYS"
4451. .5146
/gene="alr0598"
4451. .5146

/gene="alr0598"
/notes="ORF_ID:alr0598
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72556.1"
/db_xref="GI:17129944"
/translation="MVRTMRPLLSLTSLVLLQLQFPVTVQAGSENTGDVASTAIQOV
IAANLMTNSADGNFYPERLMNRRAELASILVKAFRLNKAEANQPKSITVADVPRSHWA
YQDIQTVLKTDMKGYRNLFPNQRVTRAELAIFAQAYGVQFSDDVNVNEILAPYP
DANSIPNWARRAIATVITEGFINTDQNNITPLRPMTRGDMAYLLSKYLQKQKQPD
PVPPTTDSPLP"
5596. .7503
/gene="alr0599"
5596. .7503
/gene="alr0599"
/notes="ORF_ID:alr0599"
/codon_start=1
/transl_table=11
/product="1-deoxy-xylulose 5-phosphate synthase"
/protein_id="BAB72557.1"
/db_xref="GI:17129945"
/translation="MHLSEITHPNQLHGLSIRLOQIQIARQIRDKHLQTVABFGGHLGP
GLGVVELTGLYQTLDLDRDKVINDVGHQAYEPHKLITGRYSDFHTLRQKDGITAGYLRK
GENKFDHFGAGHASTISAALMALARDMNGEKKAVAVIGDALTGMALEAINHAG
HLPTKLLVLNDNEMSTSPNVGAI PVLNKNWELSPVQFIKDNFEQKPHIPVGES
LSPELGRICEKMKRLAVPKVCANFEEGLFTTNGPVDGHNLEELIATFOQAHQIAGPVL
VHVATIKGKGYELAEKQOVGHAQTFFNLITTKAIPSNKPKPPAYAKVFSHTLVKLA
NQPKIIGITAAAMATGTDLQAKLPNQYIDVGAEOHVAHTLAAGLATEGMRPVAIY
STFLQRAQDYIIHDVCTIQNLFPVFCLODRAGIVGSDGPTHQGMYDIAYLRCIPNIVMA
PKDAEMQRMVVTGIEHTTGPIAMRFRGNGYGVPLMEGWEPIEIGKEILRNGDDV
LIIGYQWVPSMQAAETLSSEHJEATVINARFVKELDTLIVPLAOKIGRVTLEEG
CVMGFGSAVAELADLDVVVPVKRIGIPDVLVHAHPDESKAELGLTSRQIAERVLQ
AYFQKQSPVS"
7793. .8281
/gene="alr0600"
7793. .8281
/gene="alr0600"
/notes="ORF_ID:alr0600
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72558.1"
/db_xref="GI:17129946"
/translation="MADELKVKIECFILSKHMKLNFMMKKLSNLRWTATTVFCL
SAIAFVWQGAFFDATSAMDNSTYLIATSDATSLRNKSKFIRDTKNVKAARTNA
SRVDDATDNGSTAERKANKDASRIQKGAEDARTETAVDNNAVERTIDNIKGAF
NN"
complement(8399. .9469)
/gene="all0601"
complement(8399. .9469)
/gene="all0601"
/notes="ORF_ID:all0601
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72559.1"
/db_xref="GI:17129947"
/translation="MSIVFRDLLKKVSGNHTGESLNRVEAATATKMFGLGEATPAQI
GAFLIAHRIKRPTEGIELGMDLDAFDELGPKLOPISSKRPVIVFGIPYDGRTRTAPISP
VTALLMATVQPVVIHGGDRLPKYGLPLTEIMQGLGVNWTLSLEKTOEVFAQTIG
FVYTPKHEPLTNSIWEYEDOLGKRPPLATMELIWCYPAGDAHVIAGYVHPPEGMFOV
ALELRGVSKFTLVKLGSGCDLPRDRTAISLSASFELEKRVFLSTHEYIGFTTKNPL
GSTBELLTIQDVLGKSGCDLQALWNGGFYLRWSGICADMRSGIAKAEELNSGLV
AAKLQELIQVNVSVTENLSIT"
complement(9680. .10618)
/gene="ntcB"
complement(9680. .10618)
/notes="ORF_ID:all0602
ntcB gene product"
/codon_start=1
/transl_table=11
```

```
/product="nitrogen assimilation transcriptional activator"
/protein_id="BAB72560.1"
/db_xref="GI:17129948"
/translation="MRLEQLOFLAIAITSGFQKAAKCGVQSTISRQIQSLSEADLG
LELHRTNHSKLITGELRLPRARKICLEWETATQELGDLJAGKQPELCIAAIIHSLG
SYLPPVQKFRDYPVQVLRVTSIGSDRALKVLKGLVDLAI VNNRFLATTGRDMVVE
VLYDEPILLLTAANHPLAAVERVPSWSELVRYPOVFKDGYGMORLVOEKPERLEATIQ
AAWENTIDARFQVVRQELIALIPSSALVEARLDPTLAVRPLANSALPENSGLTRRV
NMVTIDRLQIPPIKHFQWLVRENIPIIVERQSRAS"
/gene
CDS
complement(10861..11166)
/gene="all10603"
complement(10861..11166)
/gene="all10603"
/note="ORF_ID:all10603
putative transcriptional regulator"
/codon_start=1
/transl_table=11
/protein_id="BAB72561.1"
/db_xref="GI:17129949"
/translation="WKTKPFDLRKMTPEQQAASEMQANLVILKLTSELRESLQGT
QGDVAKNGVVQSALSKEIHHDDDIQISTLSRYIKSLGSLTIIACFPDQEVVISQFD"
/gene
CDS
complement(11163..11522)
/gene="all10604"
complement(11163..11522)
/gene="all10604"

Query Match      21.6%; Score 351.4; DB 1; Length 339650;
Beat Local Similarity 57.9%; Pred. No. 3.4e-88;
Matches 647; Conservative 0; Mismatches 461; Indels 9; Gaps 1;

QY 209 CTGCTCGCTAGAGGTGAAGGCATCTGGAAACACAGTTTGGAAACTTTCAGGTTGCAA 268
DB 232378 CGCGTCTCTTTTCTCAAGATCATGGGCATACCTTTTGGACATCTGTTTCGATCACTA 232319

QY 269 CCTATGTTGAATCTCATGGGGTGTGTGTGTGTATCAGTGGTTGTCCACCTAGAA 328
DB 232318 CTTTGTGGCGAGTCACACGCTGGAGCGTGGGGTGTGTGATTGACGGCTGTCTCGCTAC 232259

QY 329 TTCACCTACTGAGCGACACCTACAAGTTGAACTCGATCGAAGACGGCCCGGACAGCA 388
DB 232258 TCGAAATTTCTCTGAGGAAATTCACCTGGAGTTAGACCGCAGACGCCCGGACAGATA 232199

QY 369 GAATAACCTCCACAAGGAGGAGCTGATACATGCAAAATTTCTGTCGGGACACATGAAG 448
DB 232198 AGATTACAAAGCCTCGTAAGGAGCAGACACCTGTGAGATTTGTCTGGGTATATGAAG 232139

QY 449 GGGTGACTACTGGAAAGCGCAATTTCTTGTATTGTCTCCAAACACAGATCAAAATAGCAGTG 508
DB 232138 GCAAAACTCTGGAAAGCGCAATATCAATTTGGTGGAAATAAGAACTACTCGTCCCAAG 232079

QY 509 ATCACCCTGAAATAGCCAAATGTGTACCGACCTTCTCATGCGAGACGCAACTTATGACTTCA 568
DB 232078 ATTATGATGAGATGGCGCAAAAGTATCGTCTTCCATGCTGATGCAACTTATGATGCTA 232019

QY 569 AGTACGGTGTAGACTGTACAGGAGGTGGAGGTCTCTCGGCGAGAAAACCGTTGGAA 628
DB 232018 AATATGGGATTCGTAAATTTGGCAAGGTGGCGGTAGATCTCTAGCGGGAGACAATTTGTA 231959

QY 629 GGGTGGCTGCGAGGGCCCTCCCAAGAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCT 688
DB 231958 GAGTAGCCGACGTTGGATGCTATAAATAATTTCTCGGC-----AAGTTGCTAATG 231908

QY 689 TGTCTGTTGTTTCCAAAGTGCAATAGGTTGTGCTCCAGAGACGCGGTTGATATGGGT 748
DB 231907 TAGAAGTTATAGGCTACGTCAAACGCTATCAAGAGCTTGGAAAGGTGTTGTAGATCCCAATA 231848

QY 749 CTGTAACTTTGAAACAGATAGAGACCAATCGTTAGATGCTCATCCAGAGTACGAG 808
DB 231847 CCGTCAACCTTAGCAACAGTGAAGCAATATCGTTGCTGTCTGATGGCGAATTTGGCAG 231788

QY 809 AGAAGATGATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTTCGGTGAAGTGA 868
DB 231787 ACCGCATGATTGAATTAATCGAAACAAACAGGTAGACAAGGTGATTCTTATCGCGGTGATG 231728
```

```
QY 869 TCACATGCGTCTGTAGAAAGCTTCTCTCGCGGGCTCGGTTCTCTGTGTTCGACAAAGCTCG 928
DB 231727 TGGAAATCGTAGCGCGGAATGTTCTCTAAAGGCTTGGGTGAACCTGTATTCGATAAGTTGG 231668

QY 929 AATCCGAATCTGGCAAAAGCTATGCTTTCTATTCTCTCGAGCAACCGGTTTCGAGATTGGCA 988
DB 231667 AAGCAGATATCGCAAAAGCTGTGATGTCCTTCTTACTCTGTAGCAAGGTTTGAATCGGTT 231608

QY 989 GCGGATTCGCGGAGACCGACTTTCACAGGAAGTGAGCATAATAGTATGATTTTATATGATA 1048
DB 231607 CAGGTTTTCGCGGACTCTCTCTAAACAGGATTTGAGCATAACGACGAATATATATTATGATG 231548

QY 1049 AGGCTGGAAGTGTACGACACGAGCACTAATCTCGCTCGGTTGTGTGTCAGGGAGGATATCGA 1108
DB 231547 AAAACGGTGAGATTCGCACACAGTAACCTTAAACCGTTCGCGTGGGATTTCAAGGCGGTATTC 231488

QY 1109 ATGTTGAGATTTGTGCATCTTCAAAAGTTGCTTTTAAAGCGCACACCATCTATCGGGGTGAAC 1168
DB 231487 ACGGAGAGATATCATTTTTCGAGTTGCAATTTAAACCCACAGCACAGTATTAGAAAAGAGC 231428

QY 1169 AGAACACCGTGTCAAGGAGCGCTCAGAAACGTTGAGCTTCTTAGCAAGAGGCGCATGACC 1228
DB 231427 AAAAAACAGTTACTCTGTGAAGGCGAAGAAACCTATTAGCGCCAAAGGTAGACATGATC 231368

QY 1229 CATGCGTCCGCTCTGAGCTGTTCTCTGTGTGGAATCCATGCGCGCTTGGTCTCTCATGG 1288
DB 231367 CCTGTGTTCTTACCCCGTGCAGTCCCAGATGGTAGAAGCAATGGTGGCGTGTGTTACTGTGCG 231308

QY 1289 ACCAGCTGATGGGCGACGCTGCGTCAGTGCAGTGCAGATGTT 1325
DB 231307 ATCATTTGTTACGCCATCATGGACAGTGTAAAGGTGCT 231271

RESULT 15
AP006487 1/c
WPCOMMENT
Sequence split into 6 fragments LOCUS AP006487 Accession AP006487
Fragment Name Begin End
AP006487_0 1 110000
AP006487_1 100001 210000
AP006487_2 200001 310000
AP006487_3 300001 410000
AP006487_4 400001 510000
AP006487_5 500001 523414
Continuation (2 of 6) of AP006487 from base 100001 (AP006487 Cyanidioschyzon merolae str.

Query Match      21.5%; Score 349; DB 2; Length 110000;
Beat Local Similarity 57.5%; Pred. No. 1.3e-87;
Matches 650; Conservative 0; Mismatches 475; Indels 6; Gaps 1;

QY 211 GTCGCTCTAGAGGTGAAGGCATCTGGAAACACGTTTGGAAACTTACTTTAGGTTGCAACC 270
DB 83966 GCTCGTGAAGTTGTTCAATGGGTCTCGGCAACTCTTTTGGGAGGATATTTTCGAAATAACGACT 83907

QY 271 TATGGTCAATCTCATGGGGTGGTGGTTGGTTGTTATCATGTGTTCTCCACCTAGAAAT 330
DB 83906 TTTGGTGAATCCCAACGGTGGAGTCGGAGTTGTCAATGATGTTGCTGTCACCAACGGCTG 83847

QY 331 CCACCTCACTGAGGCGACACCTACAAAGTTGAACTCGATCGAAGACGGCCCGGACAGACAGAGA 390
DB 83846 CCACCTTACTATAGAAGACGTTTCAGATTGAGTTAAACAGGCGGTAGACCGGGACAGCTAGG 83787

QY 391 ATAACTCCACAAGGAAGGAGACTGTGATACATGCAAAAATTTCTGTGAGGAGACATATGAAGGG 450
DB 83786 CTCACGACACACACGGAACGAGCAGATGAAGCGGAGATCCTCTCTGCGCTGCTCGACGGT 83727

QY 451 GTGACTACTGGAACCGCAATTTCTTGTATTGTCTCCCAACACAGATCAAAATAGGACAGTGAT 510
DB 83726 GTCACTATGGGTACACCTATAGCTATCTGTTGTTGTAACACACAGCAACCAAGGCTCTCAGAT 83667

QY 511 CACCGTGAATATAGCCAAATGTGTACCGACTTCTCATGACAGACGCAACTTATATGACTTCAAG 570
```

```

Db 83666 TATGAGGCAATGAAGTAAGCTATCGACCATCGCAGCGGATGCAACCTATGATGCAAAAG 83607
Qy 571 TAAGGTGTTAGAGCTGTACAGAGAGGTGGAGGTCTCGGGCAGAAAACCGTTGGAAGG 630
Db 83606 TACGGCATACGGCGCGTCTCGGTGGAGCGTTCAATAGCTCGGGAACAATAGGACGT 83547
Qy 631 GTGGCTGCAAGGGCCCTCCCAAGAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTG 690
Db 83546 GTTGTCTGACAGGCCCATTCGCAAAAAGATCTCTCGGACATATGCGGAACACAAATCTT 83487
Qy 691 TCGTTTGTTCCTCAAGTGCATCAGGTTGTGCTCCAGAAAGACGCGGTTGATATGGGTCT 750
Db 83486 GCGTAGTGAAGAGCGTGAAGCATATCGAG-----GCAGCGAATATTGATCCCAACGTG 83433
Qy 751 GTAACCTTTGGAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACCGAGAG 810
Db 83432 GTTTCGTACGAGCAATCGAATCGAACATCGTTCTGCTGCCCTGATTCAGAGGACGATCT 83373
Qy 811 AAGATGATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTGATC 870
Db 83372 CAAATGATCCAGCTTATCGATGAATTCGCCGAAACAGGTAACAGTGTGCGCGGTGTGTT 83313
Qy 871 ACATGCTGCTAGAAAAGCTTCCTCGCGGCTCGGTTCTCTGTGTTCGACAAAGCTCGAA 930
Db 83312 GAATGCGTTGCAAGGAACGTCCTCGCGGCTAGGATCACCGGTGTTTGACAAGCTCGAA 83253
Qy 931 TCCGAACTGGCAAAAGCTATGCTTTCTATTCTCGGAGCAACGCGGTTGAGATTGCGCAGC 990
Db 83252 GCTGAACTCGCAAAAGCCCTCATGTGCTTCAGCATCGAAAGGCTTTGAAGTAGGCTCG 83193
Qy 991 GGATTTCGCGGACCGACTTGAACAGAGTGAAGCATATGATGAGTTTATATGATGAAG 1050
Db 83192 GGTTCGCGGGCACCTTGCTAACGGGTGTAGAGCACACGATCCCTTCTATATAGTAGC 83133
Qy 1051 GCTGGAAGTGTGAGACACGAGCTAATCGCTCGGTGTGTGTCAGGGAGGGATATCGAAT 1110
Db 83132 AATGTTAGGACCGGAGCGGAGGAGAACCGTTCCGGAGGTATTTCAAGAGGCAATCTCGAAC 83073
Qy 1111 GTTGAGATTGTGCACCTTCAAGCTTCTTTTAAAGCCGACACCATCTATCGGGGTGAACAG 1170
Db 83072 GGAGAGAACATTGTATGTGCGCTCGCTTCAAAACCGACCTCTACGATATCACGCGAACAG 83013
Qy 1171 AACACCGTGTCAAGGAGCGTCAGAACGTTGAGCTTTCTAGCAAGAGGGCGCCATGACCCA 1230
Db 83012 AATACAGTGAACGCTGCGGTGAGCAAGTAGTCTTCGCGCAAGGGCAGACACGATCCC 82953
Qy 1231 TCGGTGCGCCCTCGAGCTGTTCTCTGTGTGGAATCAATGCGCGCTTGGTCTCTATGGAAC 1290
Db 82952 TGTGTGTTGCTCGTGTGCTTCCGATGGTAGAAGCCATGTCGCGCTCGTCTCTGTGTGAT 82893
Qy 1291 CAGCTGATGCGCGCAGTGGCTCAGTGCAGAGATGTTGCGGCTCAATACCTGCA 1341
Db 82892 CACTTGCTTCAACAGATTGCCCGAGTGTGCGTCTCTCTATTTGATCCATCA 82842

```

Search completed: August 27, 2005, 02:58:23  
Job time : 6886 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	1626	100.0	1626	3	Aaz50438	Aaz50438 Corn clon
2	870.4	53.5	1635	3	Aaz50435	Aaz50435 Corn clon
3	621	38.2	966	3	Aaz50440	Aaz50440 Rice clon
4	537	33.0	1595	3	Aac40227	Aac40227 Arabidops
5	535.4	32.9	1311	12	ADN73862	Adn73862 Thale cre
6	535.4	32.9	1311	13	ADT91425	Adt91426 Arabidops
7	500.6	30.8	1015	3	Aaz50437	Aaz50437 Wheat clo
8	405.6	24.9	432	8	ABX48414	Abx48414 Bovine ES
9	360.6	22.2	1086	13	ADT45459	Adt45459 Bacterial
10	351	21.6	1086	13	ADT44117	Adt44117 Bacterial
11	346.2	21.3	1089	13	ADS47934	Ads47934 Bacterial
12	316	15.4	656	13	ACN53973	Acn53973 Cotton an
13	308.2	19.0	731	9	ACL25323	ACL25323 DNA clone
14	291.6	17.9	699	9	ACL25321	ACL25321 DNA clone
15	283.2	17.4	649	9	ACL25322	ACL25322 DNA clone
16	280.4	17.2	528	13	Aaz51837	Aaz51837 Maize pla
17	280.4	17.2	1080	13	ADT46302	Adt46302 Bacterial
18	274	16.9	1053	6	ABA91398	Ab91398 Arabidops
19	274	16.9	1053	6	ABA91392	Ab91392 Arabidops
20	262.8	16.2	1131	13	ADT47171	Adt47171 Bacterial

CC encoding chorismate synthase. This clone is isolated from cs1n cDNA  
CC library which is derived from corn silk tissue seedling. Chimeric gene  
CC comprising this cDNA operably linked to regulatory sequences is used to  
CC transform host cells to alter the level of expression of chorismate  
CC synthase. The gene and its products may be used for immunological  
CC screening of cDNA expression libraries and to create transgenic plants  
CC which may also be herbicide resistant. Synthetic peptides derived from  
CC the gene are to raise antibodies, and used in screening assays to  
CC identify inhibitors which may be useful as herbicides. Probes and primers  
CC from chorismate synthase gene are used for gene mapping and as markers  
CC for traits linked to those genes  
XX  
SQ Sequence 1626 BP; 405 A; 382 C; 439 G; 400 T; 0 U; 0 Other;

Query Match	100.08;	Score 1626;	DB 3;	Length 1626;
Best Local Similarity	100.08;	Pred. No. 0;		
Matches 1626;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GCACGAGCTCAGCTTCGTCCTCTCCGCGCGCGCGCAACTATCATCTTCATTAG	60	
DB	1	GCACGAGCTCAGCTTCGTCCTCTCCGCGCGCGCGCAACTATCATCTTCATTAG	60	
QY	61	CTCATCCAACTCTATTCGATGACGACCGTGCCCAAGCCACAGCAGGTGCGCACTCACGG	120	
DB	61	CTCATCCAACTCTATTCGATGACGACCGTGCCCAAGCCACAGCAGGTGCGCACTCACGG	120	
QY	121	GCACGGCTCGACCCCGCGGATCGCGCCCTTGCTGAGATTGCCCCAGCCCTCTCTCC	180	
DB	121	GCACGGCTCGACCCCGCGGATCGCGCCCTTGCTGAGATTGCCCCAGCCCTCTCTCC	180	
QY	181	CTCCGCTTCGCGTGACCGCTGCGCCTCTGCTCCCTAGAGGTGAAGGCATCTGGAAAC	240	
DB	181	CTCCGCTTCGCGTGACCGCTGCGCCTCTGCTCCCTAGAGGTGAAGGCATCTGGAAAC	240	
QY	241	ACGTTTGAAACTACTTTTCAGGTTGCAACCTATGCTGAACTCTATGCGGGTGGTGGT	300	
DB	241	ACGTTTGAAACTACTTTTCAGGTTGCAACCTATGCTGAACTCTATGCGGGTGGTGGT	300	
QY	301	TGTGTTACAGTGGTTGTCACCTAGAAATTCACCTACCTAGGCGAGCCTACAGTTGAA	360	
DB	301	TGTGTTACAGTGGTTGTCACCTAGAAATTCACCTACCTAGGCGAGCCTACAGTTGAA	360	
QY	361	CTCGATCGAAGACGGCCGCGACAGAGCAGAAATACCTCCACAGGAAGAGACTGATACA	420	
DB	361	CTCGATCGAAGACGGCCGCGACAGAGCAGAAATACCTCCACAGGAAGAGACTGATACA	420	
QY	421	TGCAAAATTCGTGTCAGGACACATGAAGGGTGACTACTGGAACGCCAAATCTTTGTTAT	480	
DB	421	TGCAAAATTCGTGTCAGGACACATGAAGGGTGACTACTGGAACGCCAAATCTTTGTTAT	480	
QY	481	GTCCCAACACAGATCAAAATAGGCGATGATCACCGTGAATAGCCAATGTGTACCGACCT	540	
DB	481	GTCCCAACACAGATCAAAATAGGCGATGATCACCGTGAATAGCCAATGTGTACCGACCT	540	
QY	541	TCTCATGACAGACGCAACTTATGACTTCAAGTACGCTGTAGAGCTGTACAGGAGGTGGG	600	
DB	541	TCTCATGACAGACGCAACTTATGACTTCAAGTACGCTGTAGAGCTGTACAGGAGGTGGG	600	
QY	601	AGGTCCTCGGCGAGAAAACCGTTGGAAAGGGTGGCTGCGAGGGCCCTCCCAAGAAAAT	660	
DB	601	AGGTCCTCGGCGAGAAAACCGTTGGAAAGGGTGGCTGCGAGGGCCCTCCCAAGAAAAT	660	
QY	661	CTTAAGCTCAAAATGGGATTAGAGATCTGTGCTGTTGTTTCCAAAGTGCATCAGTTGTG	720	
DB	661	CTTAAGCTCAAAATGGGATTAGAGATCTGTGCTGTTGTTTCCAAAGTGCATCAGTTGTG	720	
QY	721	CTCCAGAGACGCGGTTGATTATGGGCTGTGTAACCTTTGGAAACAGATAGAGCAACATC	780	
DB	721	CTCCAGAGACGCGGTTGATTATGGGCTGTGTAACCTTTGGAAACAGATAGAGCAACATC	780	
QY	781	GTTAGATGTCCTGATCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGAGTACGA	840	
DB	781	GTTAGATGTCCTGATCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGAGTACGA	840	

QY	841	GTTTCGAGGGATTTCGGTTCGGTGGAGTGATCAGATGCGTCCGTAGAAACGTTCTTCGCGGG	900	
DB	841	GTTTCGAGGGATTTCGGTTCGGTGGAGTGATCAGATGCGTCCGTAGAAACGTTCTTCGCGGG	900	
QY	901	CTCGGTTCTCTGTTGTTCCGAACTCCGAACTGGCAAAAGCTATGCTTTCTATT	960	
DB	901	CTCGGTTCTCTGTTGTTCCGAACTCCGAACTGGCAAAAGCTATGCTTTCTATT	960	
QY	961	CTTCGAGAACCGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTGACAGGAAGT	1020	
DB	961	CTTCGAGAACCGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTGACAGGAAGT	1020	
QY	1021	GAGCAATATGATGAGTTTTATATGGAATAAGGCTGGAAAGTGTGAGGACACGAGCTAAATCGC	1080	
DB	1021	GAGCAATATGATGAGTTTTATATGGAATAAGGCTGGAAAGTGTGAGGACACGAGCTAAATCGC	1080	
QY	1081	TCGGGTGCTGTGACAGGAGGAGATATCGAATGTTTGAGATTGTGCACATTTCAAAGTTGCTTTT	1140	
DB	1081	TCGGGTGCTGTGACAGGAGGAGATATCGAATGTTTGAGATTGTGCACATTTCAAAGTTGCTTTT	1140	
QY	1141	AAGCCGACACATCTATTCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAAGACGTT	1200	
DB	1141	AAGCCGACACATCTATTCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAAGACGTT	1200	
QY	1201	GAGCTTCTAGCAAGAGGCGCCATGACCCATGCGTCCGCCCTCGAGCTGTTCTCTGCTGTG	1260	
DB	1201	GAGCTTCTAGCAAGAGGCGCCATGACCCATGCGTCCGCCCTCGAGCTGTTCTCTGCTGTG	1260	
QY	1261	GAATCCATGCGCGTTCCTCTCATGGAACAGCTGATGGCGCACGCTGGCTCAGTGGCAG	1320	
DB	1261	GAATCCATGCGCGTTCCTCTCATGGAACAGCTGATGGCGCACGCTGGCTCAGTGGCAG	1320	
QY	1321	ATGTTTCGCGCTCAATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCGAGC	1380	
DB	1321	ATGTTTCGCGCTCAATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCGAGC	1380	
QY	1381	ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT	1440	
DB	1381	ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT	1440	
QY	1441	GGCTTGGTTTGTCTCTTTCACCGTAGTTTTTTTTTTTTTTTTTCCCGCAAGTGTGATG	1500	
DB	1441	GGCTTGGTTTGTCTCTTTCACCGTAGTTTTTTTTTTTTTTTTTCCCGCAAGTGTGATG	1500	
QY	1501	CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGCAATTTGTACAGCTTTCAATATGTA	1560	
DB	1501	CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGCAATTTGTACAGCTTTCAATATGTA	1560	
QY	1561	TCTACTTCGAAGATGATGCAATTTTATAGATGTTGGCTTGTGAAGACAAAAA	1620	
DB	1561	TCTACTTCGAAGATGATGCAATTTTATAGATGTTGGCTTGTGAAGACAAAAA	1620	
QY	1621	AAAAAA 1626		
DB	1621	AAAAAA 1626		

RESULT 2

AAZ50435	ID	AAZ50435	standard; cDNA; 1635 BP.
XX	XX	AAZ50435;	
AC	AC	AAZ50435;	
XX	XX	18-MAY-2000	(first entry)
DT	DT	Corn clone Chpc24.pk0002.h1.fis	encoding chorismate synthase.
XX	XX	Corn cDNA clone chpc24.pk0002.h1.fis;	chorismate synthase;
DE	DE	immunological screening;	herbicide resistance; antibody; gene mapping;
KW	KW	corn; 88.	
KW	XX		
OS	OS	Zea mays.	



	Key	Location/Qualifiers	
FH	CDS	52..1374 /*tag= a /product= "Chorismate synthase"	
XX			
PW		WO200005353-A2.	
XX			
PD		03-FEB-2000.	
XX			
PF		20-JUL-1999; 99WO-US016353.	
XX			
PR		21-JUL-1998; 98US-0093611P.	
XX		(DUPO ) DU PONT DE NEMOURS & CO E I.	
PA		Cahoon RE, Falco SC, Pember SO;	
PI		WFI; 2000-182687/16.	
DR		P-PSDB; AAY44890.	
XX			
PT		New chorismate synthase polypeptides used to alter the level of the enzyme and thus the level of aromatic to non aromatic amino acids in transformed plants.	
PT			
XX			
PS		Claim 3; Page 28; 39pp; English.	
CC		The present sequence is corn cDNA clone designated chpc24.pk0002.hi.fis encoding chorismate synthase. This clone is isolated from chpc24 cDNA library which is derived from corn 8 day old shoot treated 24 hours with herbicide. Chimeric gene comprising this cDNA operably linked to regulatory sequences is used to transform host cells to alter the level of expression of chorismate synthase. The gene and its products may be used for immunological screening of cDNA expression libraries and to create transgenic plants which may also be herbicide resistant. Synthetic peptides derived from the gene are to raise antibodies, and used in screening assays to identify inhibitors which may be useful as herbicides. Probes and primers from chorismate synthase gene are used for gene mapping and as markers for traits linked to those genes	
XX			
SQ		Sequence 1635 BP; 441 A; 367 C; 422 G; 405 T; 0 U; 0 Other;	
	Query Match	53.5%; Score 870.4; DB 3; Length 1635;	
	Best Local Similarity	80.5%; Pred. NO. 1.6e-240;	
	Matches 1031; Conservative	0; Mismatches 246; Indels 3; Gaps 1	
QY	83	CGACGCTGCCAGCACCAGCAGGTGGCGCACTCACGGGCAGCGTCGCACCCCCGGGA 142	
Dd	59	CGCCCGTGTGCGACCGCGGTGTCGCGAGGGGTCCACACGGTTTCTCCCOCGGCGGA 118	
QY	143	TCDGCGCTTGCTGAGATTGCCCGACGCTCTCTCCCTCCGTTCCGCGTGACCGCT 202	
Dd	119	TADGCGCGCTCCCGAGTCCGCC---CCAGTCCCTCGGGTTATCCGTGCGCGCGCTC 175	
QY	203	GCGCACTGCTCGCTGAGGTGAAGGCATCTGGAAAACACGTTTGGAAACTACTTTCAGG 262	
Dd	176	GCGCGCGCTCCAGCCTTAGGTGAAGGCATCAGGAAATGTGTTCGGGAATACTTTCAGG 235	
QY	263	TTGCAACTATGGTGAATCTCATGGGGTGCTGTGGTTGTGTTATCAGTGGTTGTTCCAC 322	
Dd	236	TTGCAACTATGGGGAATCCCATGGAGGGGTGTGGTTGGTTGTTATCAGTGGTTGCCAC 295	
QY	323	CTAAGAATCCACTCAGCTGAGGCAGACCTTACAAGTTGAACCTCGATCGAAGACGGCCGGAC 382	
Dd	296	CCAGNATTCCTCTCAGTGAGGCAGACATGCAAGTAGAAGTAAGTGNATAGAGAGCTCCGGTTC 355	
QY	383	AGAGCAGATAAACCTCCACAAGGAGGAGCTGATACATGC AAAATCTGTTCAGGAGCAC 442	
Dd	356	AAGTAGAATTACAACCCCAAGAAGGAGACTGATACATGC AAAATCTATCAGGAGCAC 415	
QY	443	ATGAGGGGTGACTACTTGGAACGCCAATCTTGTGTTATTTGTCCCAAACACAGATCAATAG 502	
Dd	416	ATGATGGGATGACTACTTGTTGTTACCAATTCACAGCTTTTGTGCCAAAACACAGATCAAGGG 475	

immunological screening; herbicide resistance; antibody; gene mapping;  
rice; ss.

Oryza sativa.

Key Location/Qualifiers  
FH 45..964  
CDS /\*tag= a  
FT /product= "Chorismate synthase"  
FT /note= "no stop codon given"

WO200005353-A2.

03-FEB-2000.

20-JUL-1999; 99WO-US016353.

21-JUL-1998; 98US-0093611P.

(DUPO ) DU PONT DE NEMOURS & CO E I.

Cahoon RE, Falco SC, Pember SO;

WPI; 2000-182687/16.

P-PSDB; AAY44895.

New chorismate synthase polypeptides used to alter the level of the  
enzyme and thus the level of aromatic to non aromatic amino acids in  
transformed plants.

Claim 3; Page 35; 39pp; English.

The present sequence is rice cDNA clone designated rl872.pk0029.98  
encoding chorismate synthase. This rice is isolated from rl872 cDNA  
library which is derived from rice leaf 15 days after germination, 72  
hours after infection of strain Magaporthe grisea 4360-R-67. Chimeric  
gene comprising this cDNA operably linked to regulatory sequences is used  
to transform host cells to alter the level of expression of chorismate  
synthase. The gene and its products may be used for immunological  
screening of cDNA expression libraries and to create transgenic plants  
which may also be herbicide resistant. Synthetic peptides derived from  
the gene are to raise antibodies, and used in screening assays to  
identify inhibitors which may be useful as herbicides. Probes and primers  
from chorismate synthase gene are used for gene mapping and as markers  
for traits linked to those genes

Sequence 966 BP; 251 A; 234 C; 264 G; 217 T; 0 U; 0 Other;

Query Match 38.2%; Score 621; DB 3; Length 966;  
Best Local Similarity 80.4%; Pred No. 1.6e-168;  
Matches 742; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

75 TCCGATGACGACCGTGCACCAAGCCACAGCAGGTGGCGCACTCAGGGGACCGCTGCACC 134  
41 TCCGATGCGCGGCCAACAGTGTCTGCAGCGGTGGCGCGCTCTCCCTCCCGCGGGCGG 100  
135 CCGCGCGATCGCGCGCTTGTGGAGTTTGGCCAGCCTCTCTCCCTCCCGCTTCGCGGT 194  
101 CGCGGGGTTCGCGCGCTTCCCGGAGTCCGCCCCGCT-----TCCCTCCGCTTCTCCGT 154  
195 GCACCGCTGCGCGCACTGTCTCGCTAGAGGTGAAGGCATCTGMAACAGGTTTGGAAACTA 254  
155 CGGCGCGCGCGCGCGCTCTCGCTAGAGGTGAAGCGCTCTGCAATGTATTGGGACTA 214  
255 CTTTCAGGTTGCAACTATGTGTGAATCTCATGGGGGTGGTGTGGTTGTGTATCAGTGG 314  
215 CTTTCAGGTTGCAACTATGTGTGAATCTCATGGGGGTGGTGTGGTTGTGTATCAGTGG 274  
315 TTGTCCACTAGAAATTCCTCACTCAGCAGCACCTACAAGTTGAATCTCAGTCAAGACG 374  
275 ATGCCACCCAGAAATCCCACTTACTGAAAGCAGATATGCAAGTAGAACTCGACCGGAGACG 334  
375 GCCCGGACAGACAGAAATAAACCCTCCCAAGGAAGGAGACTGATACATGCAAAATTTCTGTC 434

Db 335 GCCAGGCAGCAGCAGATACACCCCAAGAAAGGAGACTGACACTTGCATAATTCCTTC 394  
Qy 435 AGGACACATGAAGGGGTGACTACTGGAAAGCCAAATCTCTGTTATTTGTCTCCAAACAGAGA 494  
Db 395 AGGACACATGAAGGAATGACCACTGGGACACCAATTCATGTTTGTCTCCGAAACAGAGA 454  
Qy 495 TCAATAGGCAGTGATCACCGTGAATAGCAATGTATGTCGACCTTCTCATGACAGACGC 554  
Db 455 TCAGAGAGGGGGTGAATTACAGTGAATGGCTTAAGGCTTCAGACCTTCACATGCAAGATGC 514  
Qy 555 AACTTATGACTTCAAGTACCGGTGTAGAGCTGTACAGGAGGTGGGAGGCTCCTCGGCGAG 614  
Db 515 AACTTATGACTTCAATATACCGGTGTAGAGCAGTCAGGAGGTGGAAAGATCATCAGCAAG 574  
Qy 615 AAAAAACCGTTGGAAGGGTGTGTCAGGGGCGCTCCCAAGAAATTTCTTAAGCTCAAAATG 674  
Db 575 AGAGCCATTGGAAGGGTGTGTCAGGAGCTCTTGCAGAAAGAAATTTCTTAAGCTCAAAATC 634  
Qy 675 TGGATTAGAGATCTTGTGCTGTTTTCCTCAAGGTGCATCAGGTTGTGCTCCCAAGAGACGC 734  
Db 635 TGGAGTAGAGATCTTGGCGTTTGTGTCCTCAAGGTGCATCAAGTTGTACTTACCAAGATGC 694  
Qy 735 GGTTGATTATGGGTCTCTAACTTTGGAAACAGATAGAGCAACATCTGTAGATGCTCTGA 794  
Db 695 CGTTGATTATGACACTGTAACTTGAACAGATAGAAAGCAACATCTGTAGATGCTCTGA 754  
Qy 795 TCCAGATAGCAGAGAGATGATAGCGCAATCGACAGATGACGAGTTGAGGGGATTC 854  
Db 755 TCCAGATATGCACAGAAAGATGATGATGCACTCGATAAAAGTACGAGTTAGAGTGATTC 814  
Qy 855 GGTGCGTGGAGTGATCACATGCTCGCTAGAAAGGTTCTTCGCGGGCTCGGTTCTCCTGT 914  
Db 815 GATTGGTGGTGTGTCTACATGCAATTCGCAAGAAATGTTCTTCGTGGGATGGCTCTCCTGT 874  
Qy 915 GTTCGACAAAGCTCGAATCCGAACTCGGAAAGCTATGCTTTCTATTCCTCGAGCAACGG 974  
Db 875 ATTTGACAAACTTGAGGCTGAATTTGGCGAAAGCTATGCTTTCTCTCTCTCAAGCAAGGG 934  
Qy 975 GTTCGAGATTGGCAGCGGATTCG 997  
Db 935 GTTTGAGATCGCAGTGGATTG 957

RESULT 4  
AAC40227  
ID AAC40227 standard; DNA; 1595 BP.  
XX AAC40227;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27533.  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143552P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.



XX This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreduplication, biochemistry, signal  
CC transduction, storage lipid mobilization and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is thale cress cdna  
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
CC transcription factor, given in an exemplification of the invention.  
XX  
SQ Sequence 1311 BP; 369 A; 277 C; 318 G; 347 T; 0 U; 0 Other;

Query Match 32.9%; Score 535.4; DB 12; Length 1311;

Best Local Similarity 66.9%; Pred. No. 1.1e-143;

Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 216 CCTAGAGGTGAAGGCATCTGGAACACAGTGTGGAACTACTTTCAGGTGGCAACCTATGG 275  
DB 135 CTTCCAGATACAAAGTACTGGAAGTTCATATGGGACTATTTTCGAGTTTCAACTTTGG 194  
QY 276 TGAATCTCATGGGGTGGTGTGTGTATTCAGTGTGTTCACCTAGAAATCCACT 335  
DB 195 AGATYCATATGAGGAGGAGTGTGTATCATATGATGTGTTCCTCGTATTCCTACT 254  
QY 336 CACTGAGGAGCCTCAAAAGTTGAATCGATCGAAGACGCGCCGGACAGACAGAAATAC 395  
DB 255 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGAGGCTGGTCAGACGAGGATCAC 314  
QY 396 CTCCACAGGAGGAGAGTATACATGCAAAATTCGTTCAGGACACATGAAGGGTGAC 455  
DB 315 AACTCTTAGAAGAGACTGATATCTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGAC 374  
QY 456 TAATGGAAGCCAAATCTTTGTTATTGTCTCCAAACACAGATCAAAATAGGACGATGACCG 515  
DB 375 GACAGGACACCTATCCATGTGTGTGATACAAACACAGATCAGAGGACTTGTATACG 434  
QY 516 TGAATATAGCAATGTGTACCGACCTTCTCATGACGACGCAACTTATGACTTCAAGTACGG 575  
DB 435 TGAATATGTCGTTGCTATAGACCATCGCATGCTGATGCAACTTATGACATGAATATGG 494  
QY 576 TGTATAGCTGTACAGGAGGTGGGAGTCTCGGCGCAGAAAACCGTTGGAAGGGTGGC 635  
DB 495 TGTTCAGATCAGTGCAGGGGTGGAGAGATCTTCAGCTAGAGAGACCATTTGGAAGAGTTGC 554  
QY 636 TGCAGGGGCGCTCCCAAGAAAATCTTAAGCTCAAAATGTGATAGAGATCTGTCTGTT 695  
DB 555 TCTGTGAGCTTTGGCCAGAAAATTTTGAAGCAATTTGAGGAACTGAGATTTCTGCCTA 614  
QY 696 TGTTCCTCAAGTGCATCAGGTGTGTCTCCAGAGACGCGGTTGATTTGGTCTGTAAAC 755  
DB 615 TGTCTCGCAAGTTCCACATGTTGTACTTCCAGAGAAATTTGGTAGACACAGAAATTAAC 674  
QY 756 TTTGGACAGATAGAGAGCAATCTGTATAGTGTCTGATCAGAGTATCGCAGAGAGAT 815  
DB 675 ACTCGAACAGATAGAAAATAACATTTGTTCAGATGCCCTAATCCGAGTATGCGGAAAGAT 734  
QY 816 GATAGACGAATCGACAGAGTACGAGTTCGAGGGATTCGGTGGGTGATCATCATG 875  
DB 735 GATAGCTGGATGTATGTGTGTCAGGACAAAAGGAACTCTGTGTGGGTGTGTGACCTG 794  
QY 876 CGTCTGTAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTGTTCGACAAAGCTCGAATCCGA 935

DB 795 CATTTGTCGGAATGCTCCACGTGGGCTTGGTACACCGGTTTTTCGATAAACTTTGAAGCAGA 854  
QY 936 ACTGGCAAAAGCATATGCTTTCTATTCTCTCGAGCAACGGGTTTCGAGATTGGCAGCGGATT 995  
DB 855 ACTGGCAAAAGCTTGTATGTCGCTACCTGCAACTAAGGGATTTGAGTTTGGAAAGCGGCTT 914  
QY 996 CGCCGGGACCGACTTTCAGCAGGAAGTGAGCATATGATGAGTTTATATGGAATAAGGCTGG 1055  
DB 915 TGCAGGTACCTTTTGTGACTGGTCTTGAACACATGATGATGTTCTATACGATGAAAATCG 974  
QY 1056 AAGTGTCTAGGACACGAGCATATTCGTCTGGGTGGTGTGACAGGAGGATATCGAAATGTTGA 1115  
DB 975 AAGAATATGCTTACCAAGCAACCGCATCTGGTGGAAATTCAGGAGGAGATCTCAAATGGTGA 1034  
QY 1116 GATTCTGCATCTCAAGTGTCTTTAAAGCCGACACATCTATCGGGGTGAAAACAGACAC 1175  
DB 1035 AATAATAACATGAGATGAGCTTCAAGCCAAACATCAACAATTTGGAAGGAAGCAAAATAC 1094  
QY 1176 CGTGTCAAGGGAGCGCTCAGAACGTTTGAGCTTCTAGCAAGAGGGCGCCCATGACCCATGCGT 1235  
DB 1095 GGTAAACAGACAGAGGTAGAAACCGAATGATTCGGCGTGGTCTCATGATCTCTGTGT 1154  
QY 1236 CGCCCTCTGAGCTGTCTCTGTGTGGAATTCATGCGCGGTTGGTCTCTCATGACAGCAGCT 1295  
DB 1155 TGTTCACAGAGCTGTGCCAATGTTGGAAGCAATGTTGGTCTCTAGTTCTTGTGGATCAATT 1214  
QY 1296 GATGCGGCACGTGGCTCAGTGGAGATGTTTCGGCTCAATACTGCACCTTCAGAACCC 1352  
DB 1215 GATGCGGCATACCGCACAAATGCCATTTGTTTCCAAATNAATCCAGAGTTGCGAGNAACC 1271

## RESULT 6

ADT91426

ID ADT91426 standard; DNA; 1311 BP.

XX AC ADT91426;

XX DT 16-DEC-2004 (first entry)

XX DE Arabidopsis thaliana gene 1456.

XX KW Herbicidal; plant growth; plant development; mouse-ear cress; plant;

XX KW gene; ds.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

XX CDS 1..1311

XX FT /\*tag= a

XX FT /product= "protein essential for plant growth and

XX FT development"

XX PN US2004191850-A1.

XX PD 30-SEP-2004.

XX PF 29-OCT-2003; 2003US-00696616.

XX PR 04-NOV-2002; 2002US-0423519P.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Levin JZ, Tossberg J, Zhou Q, Aux G, Mcelver J, Thomas C;

XX PI Patton DA, Frye C, Budziszewski G;

XX DR WPI; 2004-698717/68.

XX DR P-PSDB; ADT91427.

XX PT Identifying an herbicidal compound comprises combining a protein  
XX essential for growth of Arabidopsis thaliana, with a compound to be  
XX tested for the ability to bind to, or inhibit the activity of the  
XX polypeptide.

Example 9; SEQ ID NO 27; 131pp; English.

The invention relates to a method for identifying herbicidal compounds, which involves combining a protein essential for growth of *Arabidopsis thaliana* with a compound to be tested for the ability to bind to or inhibit the activity of the polypeptide. The invention also relates to nucleic acid sequences encoding proteins essential for plant growth and development. The methods, nucleic acid molecules and proteins of the invention are useful for identifying herbicidal compounds and for killing or inhibiting the growth or viability of the plant. The present sequence is the *Arabidopsis thaliana* gene encoding protein essential for plant growth and development.

Sequence 1311 BP; 369 A; 277 C; 318 G; 347 T; 0 U; 0 Other;

Query Match 32.9%; Score 535.4; DB 13; Length 1311;  
Best Local Similarity 66.9%; Pred. No. 1.1e-143;  
Matches 761; Conservative 0; Mismatches 376; Indels 0;

Qy	216	CCTAGAGGTGAAGGATCTGGAACA	CGTTTGGAAAC	TACTTTCAGGTTGCAACTATGG	275
Db	135	CTTCCAGATACAAGCTACTGGAAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTTGG		194	
Qy	276	TGAATCTCATCGGGGGTGGTGTGGTTGTGTTATCAGTGGTGTGTCCACTAGAAATTC		335	
Db	195	AGAATACATGGAGGAGGTGGTGTGTATCATTTGATGGTGTGCTCTCGTATTTCCACT		254	
Qy	336	CACGTGAGGCAGACCTTACAAGTGTAACTCGATCGAAGACGGCCCGGACAGAGCAGAAATAAC		395	
Db	255	TACTGAAATCTGATTTGGAAATTCGATTCGATAGAAAGGAGGCTGGTCAGAGCAGGATCAC		314	
Qy	396	CTCCACAAGGAAGGAGACTGATACATGCAAAATTTCTGTCAAGGACACATGAAGGGGTGAC		455	
Db	315	AACTCCTAGAAAGAGACTGATCTTGGCCGATATCGTCTGGAGTCTCTGAAGGAATGAC		374	
Qy	456	TACTGGAAACGCCAATTTCTTGTATTGTGCCAAAACACAGATCAAAATAGGCAGTGTATCCG		515	
Db	375	GACAGGAACACCTATCCATGTGTTGTATCCAAAACACAGATCAGAGAGGACTTGTATTACAG		434	
Qy	516	TGAATATGCCAATGTGTACCGACCTTCTCATGCGAGACCAACTTATGCACTTCAAGTACGG		575	
Db	435	TGAATGTGCGTTGCCATATAGACCATCGCATGCTGTATGCAACTTATGACATGAAGTATGG		494	
Qy	576	TGTTAGAGCTGTACAGGAGGTGGGAGTCTCCGGCGACGAAAAACCGTTGGAAGGGTGGC		635	
Db	495	TGTACAGATCACTGCAGGGTGCAGGAAGATCTTTCAGCTAGAGAGACCATTTGGAAGAGTTGC		554	
Qy	636	TGCAAGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTT		695	
Db	555	TCTTGGAGCTTTGGCCAAAGAAAAATTTTGAAGCAATTTTGCAGAACTTCGAGATTTCTTGCCCTA		614	
Qy	696	TGTTTCCAAAGTGCATCAGGTGTGCTCCCGAGAGACGGGTTGATTTATGGTCTCTGTAAAC		755	
Db	615	TGTTCTGGCAAGTTCAACATGTTGACTTCCAGAGAAGAAATGGTAGACCAACGAGAATTTTAAAC		674	
Qy	756	TTTGGAAACAGATAGAGACCAACTCGTTAGATGTCTTGATCCAGATCCAGAGTACGACAGAGAAGAT		815	
Db	675	ACTCGAACAGATAGAAATAACAATGTACAGATGCCCTAATCCCGAGTATGGCGGAAAAGAT		734	
Qy	816	GATAGAAGCAATTCGACAGAGTACGAGTTTCGAGGGGAATTCGGTGGAGTGTATCACATG		875	
Db	735	GATAGTGCAGATTGATGCTGTCCAGACAAAAGGGAACTCTGTTGGTGGTGTGTGACCTG		794	
Qy	876	CGTTCGCTAGAAACGTTTCTCCGGGGCTGGTTCTTCTGTGTGTTCACAAGCTCGAATCCGA		935	
Db	795	CAITGTTTCGGAATGCTCCACGTGGGCTTGGTACACCCGGTTTTCGATAAACTTGAAGCAGA		854	
Qy	936	ACTGGCAAAAGCTATGCTTTCTATTCTCTCGCAGCAACCGGGTTTCAGAGATTGGCAGCGGATT		995	
Db	855	ACTGGCAAAAGCTTGTATGTTCGCTACCTGCAACTTAAGGGATTTGAGTTTGGAAACGGCTT		914	
Qy	996	CGCCGGGACCGACTTGAACAGGAAGTGACATAAATGATGAGTTTTATATGGATAAAGGCTGG		1055	

peptides derived from the gene are to raise antibodies, and used in screening assays to identify inhibitors which may be useful as herbicides. Probes and primers from chorismate synthase gene are used for gene mapping and as markers for traits linked to those genes

[illegible]

RESULT 8  
ABX48414  
ID ABX48414 standard; cDNA; 432 BP.  
XX  
AC ABX48414;  
XX  
DT 21-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #13579.

[illegible]

Bovine; ss; EST; expressed sequence tag; lactation; LMFD; muscle deposition; fat deposition; genome mapping; gene identification; gene analysis; cattle breeding.

**Bob Taurus.**

US2002137139-A1.

26-SEP-2002

24-SEP-2001

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

11-JAN-2000; 2000US-00480902.

(BYAT/) BYATT J C.

(TAON/) TAO N.

J. M. NEWMAN / (NYFPA)

BYATT JC, Machiata

WPI; 2003-110599/10.

New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.

Claim 2; SEQ ID NO 13579; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ARX34836-ARX49947, or complements of them. Also included are:

- (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non- translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
- (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site.

CC seqdata.uspto.gov/sequence.html?DocID=20020137139

SQ Sequence 432 BP; 91 A; 102 C; 125 G; 114 T; 0 U; 0 Other;

Query Match	24.9%;	Score 405.6;	DB 8;	Length 432;
Best Local Similarity	97.7%;	Pred. No. 2e-106;		

QY  
...  
db

Qy	1143	GC	GAC	CA	CC	AT	CT	AT	CG	GG	GT	GA	AA	CA	GA	AC	CG	TG	CA	AG	GC	GT	CA	GA	CG	T	TGA	1202	
Db	61	GC	GAC	CA	CC	AT	CT	AT	CG	GG	GT	GA	AA	CA	GA	AC	CG	TG	CA	AG	GC	GT	CA	GA	CG	T	TGA	120	
Qy	1203	GC	TT	CT	AG	CA	AG	GG	GG	CC	AT	CA	CC	AT	GC	GT	CG	CC	CC	CT	CG	AG	CT	TT	CT	GT	GG	TGA	1362



|||||  
Db 121 GCTTCTGCAAGAGGCGCCATGACCCATGGTGGTCCCTCGAGCTGTTCTGTGGTGA 180  
Qy 1263 ATCCATGCGCGTTGGTCTCTCATGGACCAGCTGATGGCGACGTTGGCTCAGTGGAGAT 1322  
Db 181 ATCCATGCGCGTTGGTCTCTCGTGACCCAGCTGATGGCGACGTTGGCTCCAGTGGAGAT 240  
Qy 1323 GTTCGGCTCAATATGCACTTCAAGAACCAAGTTGGCTCTTCTTAGCAGAGGCGAGAGCAC 1382  
Db 241 GTTCGGCTCAATGCTGCACTTCAAGAACCAAGTTGGCTCTTCTTAGCAGAGGCGAGAGCAC 300  
Qy 1383 ACCTGATGAGCTCGGCGCAAAATTTATCATATTAATCATAGTAATAGTAGCTCAAGCGTGG 1442  
Db 301 ACCTGATGAGCTCGGCGCAAAATTTATCATATTAATCATAGTAATAGTAGCTCAAGCGTGG 360  
Qy 1443 CTTGGTTTGTGCTCTTGTGACCGTAGTGTGTTTTTTTTT-TTCCCGCAAGTGTGATGC 1501  
Db 361 CTTGGTTTGTGCTCTTGTGACCGTAGTGTGTTTTTTTTTCCCGCGCAAGTGTGATGC 420  
Qy 1502 GATGAAGTGAAT 1513  
Db 421 GATGAAGTGAAT 432

## RESULT 9

ADT45459  
ID ADT45459 standard; cDNA; 1086 BP.  
AC ADT45459;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polynucleotide #20210.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polynucleotide; gene; ss.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 43897; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a

transformed plant having an improved property. The plant is a crop plant  
such as maize or soybean. The method of producing a transformed plant  
having an improved property comprises transforming a plant with the  
recombinant DNA construct and growing the transformed plant, where the  
polynucleotide or polypeptide is useful for improving plant properties.  
The recombinant DNA construct is useful for producing plants with  
improved plant properties, e.g. improved cold, heat or drought tolerance,  
tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
increased resistance to plant disease, better growth rate by modification  
of the cell cycle pathway with plant growth regulators, increased rate of  
homologous recombination, modified seed oil or protein yield and/or  
content, improved yield by modification of carbohydrate, nitrogen or  
phosphorus use and/or uptake, by modification of photosynthesis or by  
providing improved plant growth and development under at least one stress  
condition. This sequence represents a bacterial polynucleotide used in  
the scope of the invention. Note: The sequence data for this patent did  
not form part of the printed specification but was obtained in electronic  
format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 1086 BP; 319 A; 201 C; 290 G; 276 T; 0 U; 0 Other;

Query Match 22.2%; Score 360.6; DB 13; Length 1086;  
Best Local Similarity 58.9%; Pred. No. 3.4e-93;  
Matches 643; Conservative 0; Mismatches 439; Indels 9; Gaps 1;  
Qy 235 GGAACAACGTTTGGAACTACTTTCAGGTTGGACCTATGTTGTAATCTCATGGGGTGGT 294  
Db 4 GGCAATATTTTGGTCAATTTATTTCCGCAATAGTACTTTTGGCGAGTCTCAGCGGGGGT 63  
Qy 295 GTTGGTTGTGTATCATGTTGTGTCACCTAGATTTCCACTCCTAGGAGGACCTACAA 354  
Db 64 GTGGGGGTGTGATGTGATGTTGTCTCCACAACTAGAAATTTCCGAGAGAAGAAATTCAG 123  
Qy 355 GTTGAACTCGATCGAAGACGCGCGGACAGACAGAGAGATAACCTCCACAAGAGAGAGACT 414  
Db 124 GTAGAACTAGATAGAGCGCGCGGACAAAGTAAATTTAGCACTCCCGCCAAAGAGCT 183  
Qy 415 GATACATGCAAAATTTCTGTGAGGACACATGAAGGGGTGACTACTGGAACGCCAATTCCT 474  
Db 184 GATACCTGCGAGATTATCTCAGGAGTATTTTGAAGGCAAAACGGTGGGAACCCCTATAACG 243  
Qy 475 GTTATTTCTCCAAACACAGATCAAAATAGGAGTGTATCACCCTGAAATAGCCAATGTGTAC 534  
Db 244 ATTTTGTAGTATCAAGATATCTGTCGCCCAAGATTATGACGAGATGGACAGAAATAT 303  
Qy 535 CGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGGA 594  
Db 304 CGGCTTCTCAGCGGATGCAACCTATGATGCAAAATATGGCAATTTGGCAAGGT 363  
Qy 595 GGTGGAGAGTCTCGGCGAGAAAACCGTTGGAAGGGTGGTGCAGGGGCCCTCCCAAG 654  
Db 364 GGGGTTAGTGGTGCAGCACGCTGACAAATCGGGAGAGTAGCAGCAGGTGGCATCGCTAAA 423  
Qy 655 AAAATTTCTTAAGCTCAAAATGTGGAATAGAGATCTTGTGCTTTTGTTCCAAAGTGCATCAG 714  
Db 424 AAAATTTCT-----CCGTCAAGTTGCCAATGTCCAAATATTGTTACGTTAAGCGG 474  
Qy 715 GTTGTGCTCCAGAGACGCGGTTGATTTATGGGCTCTGTAACCTTTTGGAAACAGATAGAGAGC 774  
Db 475 ATTAAGACTTTGGAAGGTGTAGTCGATCCTTAATACTGTAACCTTTAGAACAAGTAGAAGC 534  
Qy 775 AACATCTTAGATGTTCTCTGATCCAGACTACGACAGAGAAGATGATAGACGCAATCGACAGA 834  
Db 535 AATATCGTTGCTGCTCCGATGCTGAATGTGGCGATCGCATGATTGAATTTGAATTTGAGCAA 594  
Qy 835 GTACGAGTTCGAGGGGATTCGGTGGTGGAGTGTATCATGCTCGCTCGCTAGAAAAGCTTCTCT 894  
Db 595 ATAGGTAGACAAGCGGATTTCTATCGCGGGTGTCTAGAAATGTGTGGCGGCAATGTACCG 654  
Qy 895 CGCGGGTCCGGTCTCTCTGTGTTTCACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTT 954  
Db 655 AAAGGTTTGGCGCAACCAAGTATTTGATAAATTAGAAGCTGATATCGCTAAGGGTGTGATG 714





Db 595 ACAGGTGACAAAGGTTCTTATCGCGGTGTAGTGGATCGGTAGCGGAATGTTCT 654  
Qy 895 CGCGGCTCGGTTCTCTGTGTTCGACAAAGCTCGAATCGGAATCGGAAAAGCTATGCTT 954  
Db 655 AAAGGCTTGGGTGAACCTGTATTTCGATAAGTTGGAAGCAGATATCGCCAAAGCTGTGATG 714  
Qy 955 TCTATTCCTCGAGCAACGGTTTCGAGATTGGCAGCGGATTCGCGGACCGACTTGACA 1014  
Db 715 TCTTTACCTGTAGCAAAAGGTTTGAATCGGTTTCAGGTTTTCGCGGACTCTCTAACA 774  
Qy 1015 GGAAGTGAACATATGATGAGTTTATATGATGAAGCTGGAAGTGTTCAGACACGGACT 1074  
Db 775 GGAATTGAGCATACGACGATATATATTGATGAAACGGTGAGATTGCGACAGTAAT 834  
Qy 1075 AATCGTCGGGTGTGTTCAGGGAGGATATCGAATGTGTGATGTGCACTTCAAAAGTT 1134  
Db 835 AACCGTTCCGGTGGGATTCGAAGGGGTATTGCCAACGGAGAGAAATATCAITTTGCGAGTT 894  
Qy 1135 GCTTTAAGCCGACACCATCTATCGGGGTGAAACAGACACCGGTGTCGAAGGGAGCTCAG 1194  
Db 895 GCATTTAAACCCACAGCCACGATTAAGAAAGAGCAAAACAGTTACTCGTGAAGGGGAA 954  
Qy 1195 AACGTTGAGCTTCTAGCAAGGGCGCCATGACCCATCGCTCGCCCTCGAGCTGTTCT 1254  
Db 955 GAAACCTATTAGCGCCAAAGGTAGACATGATCCCTGTGTCTACCCGTCGAGTCCG 1014  
Qy 1255 GTGTGGAATCAATGCGCGCTGTGTCTCTATGACACAGCTGATGGCGCAGTGGCTCAG 1314  
Db 1015 ATGTTAGAGCAATGTGTCGCTGTGTACTGTGCGATCATTTGTTACGCGCATCATGACAG 1074  
Qy 1315 TCGGAGATGTT 1325  
Db 1075 TGTAAGGTGCT 1085

RESULT 11  
ADS47934

ID ADS47934 standard; cDNA; 1089 BP.

XX AC ADS47934;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #2677.

XX KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polynucleotide; gene; BS.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 26364; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polynucleotide used in  
CC the scope of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1089 BP; 278 A; 245 C; 314 G; 252 T; 0 U; 0 Other;

Query Match 21.3%; Score 346.2; DB 13; Length 1089;

Best Local Similarity 58.1%; Pred. No. 5e-89;

Matches 634; Conservative 0; Mismatches 448; Indels 9; Gaps 1;

Qy 235 GGAACACGTTTGGAACTACTTTTCAGGTTGCAACCTCTGTTGATCTCTGAGGTTGTT 294  
Db 4 GGAACACGTTTGGAGTTTATTTTCGCATTACCACTTCGGTGAGTCCCATGCGGCGGC 63  
Qy 295 GTTGGTTGTGTATCAGTGGTTGTCACCTAGATAATTCACCTCAGTGGAGAGACTACAA 354  
Db 64 GTGGGGTAAATTATTTGATGTTGCCCTCCCGATTTGGAATTTCCCCAGAGGAATCCAA 123  
Qy 355 GTTGAATCGATCGAAGACGGCCGGACAGACAGACAGATAACCTCCACAGGAAGAGACT 414  
Db 124 GTAGACCTAGACCGCGCGCCAGGACAGACAGCAAAATACCACTTCGTAAGGAAGCA 183  
Qy 415 GATACATGCAAAATTTCTGTCAGGACACATGAGGGGTGACTACTGGAACGCCAATTTCT 474  
Db 184 GATCAATGCGAAATTTCTCTGCGGTGTTTCGAGGGTAAACCTCTGGGTACACCATCGCC 243  
Qy 475 GTTATTGTCCAAACACAGATCAATAGGAGTGTATCAGCTGAAATAGCCCAATGTGTAC 534  
Db 244 ATTTGTGTGAAATAGGATGCCCGTCCAGAGATTACATGAGATGCGGTGAAATAT 303  
Qy 535 CGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACCGGTGTTAGAGCTGTACAGGA 594  
Db 304 CGTCTTCCACGCGGATGCCACCTATGAAGCGAAATATGGCATTCGCAATTCGCAAGGA 363  
Qy 595 GGTGGAGGTCTTCGGCAGAAAACCGTTGGAAGGGTGGCTGCGAGGGCCCTCCCAAG 654  
Db 364 GGAGTCTGCTTCTCTGTAGGAAACCATTTGGCGGGTGGCCGAGGGCGATCGCCAAA 423  
Qy 655 AAAATTCTTAAGCTCAAAATGTGGAATTAGAGATCTTGTCTGTTTGTTCCTCAAGTGCATCAG 714  
Db 424 AAGATTCTGGCCCAAGTTTAACGGGGTGGAAATTCGGCTATGTGAAAGATATCCAGAC 483  
Qy 715 GTTGTGCTCCAGAAAGACGGGTTGATTATGGGTCTCTGTAACCTTTGGAAACAGATAGAGC 774  
Db 484 ATT-----GAGGCGACGGTGGACAGCAATACGGGTGACCTTTGGAGCAAGTGAAGC 534



Db 362 GCAACCAAGGTTTGAATTTGGGAGTGGAATTTGCAGGTACTTTCTTGACCGGTAGTGAA 421  
Qy 1024 CATATGATGAGTTTATATGATTAAGCTGGAGTGTTCAGACACGAGCTAATCGCTCG 1083  
Db 422 CATATGATGAGTTTATATGATTAAGCTGGAGTGTTCAGACACGAGCTAATCGCTCG 481  
Qy 1084 GGTGCTGTGCAGGAGGAGTATCGAATGTTGAGATTGTGCACCTTCAAAAGTTGCTTTTAAG 1143  
Db 482 GGTGGATACAGGGTGGAAATCCAAUUGGGAAATTAATAATGAGAGTACGTTTCAAG 541  
Qy 1144 CCGACACCATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAAAGTTGAG 1203  
Db 542 CCAACAGCTACAAATTTGGTAAGAAACAGACACAGTACTCGAGAAAGAG-AGATAGAA 600  
Qy 1204 CTTTAGCAAGAGGGCGGCATGACCCATGCGTGCCTCGAGCTGTTCTGTGTGT 1259  
Db 601 CTACTAGCCCGGTCGTCAATGATCCTTGTGTGTGTCTCGAGCGTCCGATGT 656

## RESULT 13

ACL25323/c  
ID ACL25323 standard; DNA; 731 BP.

XX AC  
XX AC  
XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

SQ Sequence 731 BP; 185 A; 177 C; 171 G; 195 T; 0 U; 3 Other;  
Query Match 19.0%; Score 308.2; DB 9; Length 731;  
Best Local Similarity 79.6%; Pred. No. 3.9e-78;  
Matches 364; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
Qy 905 GTTCTCCTGTTCGACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTTTCTTCTTCTG 964  
Db 731 GCTCTCCTGTTCGACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTTTCTTCTTCTG 672  
Qy 965 CGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGCACAGGAAGTGAGC 1024  
Db 671 CAAGCAAGGGTTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGCACAGGAAGTGAGC 612  
Qy 1025 ATATGATGATTTTATATGATTAAGCTGGAAGTGTTCAGACACGAGCTAATCGCTCGG 1084  
Db 611 ATAACGATGATTTCTATATGATGAGGTGGAATGTGAGAACACGAAACCAATCGCTCGG 552  
Qy 1085 GTGCTGTGCAGGGAGGGATATCGAATGTTTCAGATTGTGCACCTTCAAAAGTTGCTTTAAGC 1144  
Db 551 GCGGTGTACAGGGCGGGATATCAATGGTGAACCATATATCTTCAAAAGTAGCTTTCAAGC 492  
Qy 1145 CGACACCATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTTCAGAACGTTGAGC 1204  
Db 491 CAACAGCAACTATTTGGGAAGAGCAAAATACTGTGACAAGGGATCATGAGGATATCGAGC 432  
Qy 1205 TTCTAGCAAGAGGGCGGCATGACCCATGCGTGCCTCGCCCTCGAGCTGTTCTGTGGTGAAT 1264  
Db 431 TTCTAGCAAGGGGTGCGCATGACCCATGTTGTTCCTCGGCTGTTTCCCAATGGTGGAGA 372  
Qy 1265 CCATGGCGGTTGGTCTCTCATGACACCGTGTATGGCGCACGTGGCTCAGTGCAGATGT 1324  
Db 371 CGATGGCGGATTGGTCTCTCATGACACCGTGTATGGCGCACATGTTGCTCAATGCGAGATGT 312  
Qy 1325 TCGGCTCAATATCTGCACTTCAAGAACCAAGTTGCTC 1361  
Db 311 TCGGCTCAATCTGCACTTCAAGAACCAAGTTGCTC 275

## RESULT 14

ACL25321/c  
ID ACL25321 standard; DNA; 699 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

DNA clone originating in barley containing SNP encoding sequence #15312.

Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
gene; ss.

Hordeum vulgare; var. (cul.Haruna Nijo).

W02003057877-A1.

17-JUL-2003.

16-DEC-2002; 2002WO-IB005403.

20-DEC-2001; 2001JP-00387059.

20-DEC-2001; 2001JP-00387131.

20-DEC-2001; 2001JP-00403299.

20-DEC-2001; 2001JP-00403300.

27-SEP-2002; 2002JP-00327515.

(UYN1-) UNIV JAPAN OKAYAMA.

Sato K, Takeda K, Kohara Y;

WPI; 2003-587127/55.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Single nucleotide polymorphism sites in barley varieties and DNA  
sequences containing them for analysis and identification of barley  
varieties and production of barley transformants with desired  
characteristics.

Disclosure: SEQ ID XX; 284pp; Japanese.

The present invention relates to oligonucleotide clones originating in  
barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
(SNP). The oligonucleotides may be used for analysis of SNPs among barley  
varieties, identification of particular varieties and genotype-phenotype  
analysis, isolation of specific genes and creation of new varieties by  
transformation of barley varieties with them and production of new barley  
varieties with desired properties. The present sequence represents an  
oligonucleotide clone DNA sequence featured in the specification. The  
sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to  
standardise OS field)



Search completed: August 27, 2005, 01:03:33  
Job time : 875 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 21:25:10 ; Search time 4902 Seconds

(without alignments)  
12625.969 Million cell updates/sec

Title: US-10-677-179-7

Perfect score: 1626

Sequence: 1 gcacgagctcagcttcgtct.....caaaaaaaaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	1626	AY105171	Zea mays
2	618.2	38.0	657	BM073857	MEST118-F
3	607.2	37.3	620	BM078362	MEST118-F
4	562	34.6	1160	CK210372	FGAS02217
5	550.2	33.8	763	CB654761	OSJNEC071
6	527.6	32.4	1488	CB654761	Arabidops
7	523.6	32.2	1494	CNSOACNM	Arabidops
8	513	31.5	814	CNI44654	WOUND1_23
9	499.8	30.7	1538	CNSOACNT	Arabidops
10	498.6	30.7	723	CF303160	ABF1--01-
11	487.6	30.0	710	BU099249	WHE3305_C
12	480.4	29.5	702	BU099249	WHE3305_C
13	477.4	29.4	864	CA189902	SCCCLR1G0
14	475.6	29.2	694	CD870292	AZ02_113P
15	473	29.1	690	CA124299	SCQGLR108
16	471.2	29.0	696	BQ841547	WHE4212_D
17	469.2	28.9	667	CA119896	SCCCLR106
18	468.2	28.8	749	CA094750	SCCCLR401
19	461	28.4	812	CNI29049	RHOH1_33
20	458.8	28.2	689	CA298301	SCRFL103
21	457.8	28.2	673	CF571696	MCS016H05
22	455.2	28.0	771	CA102511	SCBGRH106
23	449.6	27.7	923	CO082566	GR_Ea47E
24	443.2	27.3	718	BU100489	WHE3353_H

25	443	27.2	639	4	BG414940	HVSMek000
26	442.2	27.2	729	5	BQ789431	WHE4161_C
27	435	26.8	605	6	CA220612	SCRFL400
28	433.4	26.7	629	1	AV832717	AV832717
29	426.4	26.2	661	6	CD899216	G174_111J
30	425.4	26.2	604	6	CA214673	SCUTFL105
31	419.6	25.8	644	1	AJ432117	AJ432117
32	419.2	25.8	579	6	CA281236	SCRFLD102
33	419	25.8	596	4	BG840749	MEST11-D0
34	411.4	25.3	683	6	CD207597	HSL_33_D1
35	409.2	25.2	489	4	BM350343	MEST264-E
36	408	25.1	697	6	CD890775	G118_115G
37	408	25.1	865	7	CO114396	GR_Eb015
38	407.6	25.1	568	2	BF065345	HV_Eb001
39	405.8	25.0	592	6	CD868036	AZ02_107N
40	398.6	24.5	835	7	CO080629	GR_Ea44D
41	398.4	24.5	666	7	CK369578	zmrbw485
42	396.2	24.4	562	1	AJ435616	AJ435616
43	395.2	24.3	604	2	BE497989	WHE0958_A
44	394.8	24.3	755	7	CK279831	EST725909
45	390.8	24.0	786	4	BG600177	EST505072

## ALIGNMENTS

RESULT 1	AY105171	Zea mays	1626 bp	linear	HTC 16-OCT-2002
LOCUS	PCO087106	mRNA sequence.			
DEFINITION	Zea mays				
ACCESSION	AY105171				
VERSION	AY105171.1	GI:21208249			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	1 (bases 1 to 1626)				
AUTHORS	Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 1626)				
AUTHORS	Coe,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	location/Qualifiers				
source	1..1626				
	/organism="Zea mays"				
	/mol_type="mRNA"				
	/db_xref="MaizeDB:634903"				
	/db_xref="taxon:4577"				
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"				
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"				
ORIGIN					
Query Match	100.0%	Score	1626;	DB 3;	Length 1626;







97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.

## PCR Primers

FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA CTA TAG)  
 BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG) .

## FEATURES

source

Location/Qualifiers  
 1. .657  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="WEST76-G12"  
 /tissue\_type="Seedling and silk"  
 /lab\_host="DH10B"  
 /clone\_lib="ISUM4-TN"  
 /note="Vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGAAGATTCGGCCGCGAGATTTTCTTTTCTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT7T3PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

## ORIGIN

Query Match 38.0%; Score 618.2; DB 4; Length 657;  
 Best Local Similarity 98.5%; Pred. No. 1.5e-168;  
 Matches 645; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 973 GGGTTTCAGATTGGCGAGCGGATTCGGCGGACCGACTTCACAGGAGTGAGCATATGAT 1032  
 DB 657 GGGTTTCAGATTGGCGAGCGGATTCGGCGGACCGACTTCACAGGAGTGAGCATATGAT 598  
 QY 1033 GAGTTTATATGATAGCTGAGCTGACGACGACGACTTAATCGCTCGGCTGGTGTG 1092  
 DB 597 GAGTTTATATGATAGCTGAGCTGACGACGACGACTTAATCGCTCGGCTGGTGTG 538  
 QY 1093 CAGGGAGGATTCGAATTTGAGATTGTGCACTTCAAAAGTTGCTTTTAAGCCACACCA 1152  
 DB 537 CAGGGAGGATTCGAATTTGAGATTGTGCACTTCAAAAGTTGCTTTTAAGCCACACCA 478  
 QY 1153 TCTATCGGGTGAACAGACACCGCTGTCAGGGAGCGTCAGACGTTGAGCTTCTAGCA 1212  
 DB 477 TCTATCGGGTGAACAGACACCGCTGTCAGGGAGCGTCAGACGTTGAGCTTCTAGCA 418  
 QY 1213 AGAGGGCCCATGACCCATGCGTCCCTCGAGCTGTTCTGCTGTTGGAATCATGCGCC 1272  
 DB 417 AGAGGGCCCATGACCCATGCGTCCCTCGAGCTGTTCTGCTGTTGGAATCATGCGCC 358  
 QY 1273 GGGTTGTCCTATGACGACGCTGATGCGGCACGTCGTCAGTTCGAGATGTTTCGCGCTC 1332  
 DB 357 GGGTTGTCCTATGACGACGCTGATGCGGCACGTCGTCAGTTCGAGATGTTTCGCGCTC 298  
 QY 1333 AATCTGCACTTCAAGAACCACTTGGCTCTTCTTACAGAGGCGAGACACCTGATGAG 1392  
 DB 297 AATCTGCACTTCAAGAACCACTTGGCTCTTCTTACAGAGGCGAGACACCTGATGAG 238  
 QY 1393 CTCGGCCCAATTTATCATTTATCATAGTATGCTCAAGCGTGGCTGGTTTTCG 1452  
 DB 237 CTCGGCCCAATTTATCATTTATCATAGTATGCTCAAGCGTGGCTGGTTTTCG 178  
 QY 1453 TTGTCCTTGACCGTAGTTTGTGTTTTTTTTTCCCGCAGTGTGATGATGAAGTGA 1512

Db 177 TTGTCCTTCCCGTAGTTTGTGTTTTTTTTTCCCGAAGTGTGATGCGATGAAGTGA 118  
 QY 1513 TAAGGCACTTGGTTTCTGTCATTTGTACAGTTTTCATATAATG-TAATCTACTTCGAA 1571  
 Db 117 TAAGGCACTTGGTTTCTGTCATTTGTACAGTTTTCATATAATGCTAATCCCTTCGAA 58  
 QY 1572 GATGATGCAATTTTATAGATGGCTTGTGAAGACAAAAAATAAAAAA 1626  
 Db 57 GATGATGCAATTTTATAGATG-GGCTTGTGAAGCAAAAAAATAAAAAA 4

## RESULT 3

BM078362/c

LOCUS

DEFINITION

BM078362

ACCESSION

BM078362

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

620 bp mRNA linear EST 14-NOV-2001  
 MEST118-F09.T3 ISUM4-TN Zea mays cDNA clone MEST118-F09 3', mRNA  
 sequence.  
 BM078362  
 EST.  
 GI:16925294  
 Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 Qiu,F., Qui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.  
 1 (bases 1 to 620)  
 Expressed Sequence Tags from B73 Maize Seedlings and Silks  
 Unpublished (2001)  
 Contact: Patrick S. Schnable  
 Schnable Laboratory  
 Iowa State University  
 G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
 Tel: 515-294-0975  
 Fax: 515-294-2299  
 Email: schnable@iastate.edu  
 Individual basecall and confidence value were assigned using the  
 Phred software,  
 (<http://depts.washington.edu/ventures/collabtr/direct/index.htm>#b  
 rt). Overall sequence quality assessment and vector trimming were  
 conducted using the Lucy software (<http://www.tigr.org/softlab/>).  
 Lucy parameters were set to ensure an overall trimmed quality of  
 97.5% or better without any vector fragments in the chosen  
 high-quality region of each sequence. Low-quality bases between the  
 poly-T and the high-quality region were replaced with N's to serve  
 as spacers.  
 PCR Primers  
 FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA  
 CTA TAG)  
 BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC  
 TAA AG)  
 Seq primer: primer T3 (ATT AAC CCT CAC TAA AG) .

## FEATURES

source

Location/Qualifiers  
 1. .620  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="MEST118-F09"  
 /tissue\_type="Seedling and silk"  
 /lab\_host="DH10B"  
 /clone\_lib="ISUM4-TN"  
 /note="Vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI;  
 ds-cDNA molecules were generated as follows. First-strand  
 cDNA was prepared from oligo-dT selected mRNA by priming  
 with a NotI oligo-dT primer (5' AACTGGAAGATTCGGCCGCGAGATTTTCTTTTCTTTT). The  
 resulting DNA:RNA hybrid was treated with RNase H and used  
 as a template for DNA PolI-catalyzed second strand  
 synthesis. After the addition of EcoRI adaptors, the  
 ds-cDNAs were digested with NotI and size-selected. The  
 resulting molecules were directionally cloned into the  
 EcoRI and NotI sites of the pT7T3PAC vector. The library

then went through one round of normalization to Cof value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

ORIGIN	Query Match	37.3%;	Score 607.2;	DB 4;	Length 620;
	Best Local Similarity	98.7%;	Pred. No. 2.4e-165;		
	Matches 612;	Conservative	0;	Mismatches 8;	Indels 0;
	Gaps	0;			
Qy	1002	GACCGACTTGACAGGAAGTGACATTAATGATGAGTTTATATGATTAAGGCTGGAAGTGT	1061		
Db	620	GACCGACTTGACAGGAAGTGACATTAATGATGAGTTTATATGATTAAGGCTGGAAGTGT	561		
Qy	1062	CAGGACCGGACTAATCGCTCGGCTGGTGTGCAGGAGGGATATCGAATGTTGATGTT	1121		
Db	560	CAGGACCGGACTAATCGCTCGGCTGGTGTGCAGGAGGGATATCGAATGTTGATGTT	501		
Qy	1122	GCACCTTCAAAAGTTGCTTTTAAGCCGACACCATCTATCGGGTGAAACAGAACACCGTGT	1181		
Db	500	GCACCTTCAAAAGTTGCTTTTAAGCCGACACCATCTATCGGGTGAAACAGAACACCGTGT	441		
Qy	1182	AAGGAGCGCTCAGAACGTTGAGTCTTAGCAAGAGGGCGCCATGACCCATCGTGC	1241		
Db	440	AAGGAGCGCTCAGAACGTTGAGTCTTAGCAAGAGGGCGCCATGACCCATCGTGC	381		
Qy	1242	TCGAGCTGCTCTGTGGTGGAAATCATGCGCGCTTGGTCTCATGGACAGCTGATGGC	1301		
Db	380	TCGAGCTGCTCTGTGGTGGAAATCATGCGCGCTTGGTCTCATGGACAGCTGATGGC	321		
Qy	1302	GCAGTGCTCAGTGCAGAGATGTTGCGCGCTCAATCTGCACTTCAAGAACAGTTGGTCT	1361		
Db	320	GCAGTGCTCAGTGCAGAGATGTTGCGCGCTCAATCTGCACTTCAAGAACAGTTGGTCT	261		
Qy	1362	TTTTCAGCAGAGGCGAGACACACTGATGAGCTCGCGCCAAATTTATCATTTATCATAG	1421		
Db	260	TTTTCAGCAGAGGCGAGACACACTGATGAGCTCGCGCCAAATTTATCATTTATCATAG	201		
Qy	1422	TAATTAAGTAGCTCAAGCGTGGCTGTTGCTCTCTCTGCTGACCGTAGTTTCTTTT	1481		
Db	200	TAATTAAGTAGCTCAAGCGTGGCTGTTGCTCTCTCTGCTGACCGTAGTTTCTTTT	141		
Qy	1482	TTTTCCCGCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTCTGTCGATTTGTA	1541		
Db	140	TTTTCCCGCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTCTGTCGATTTGTA	81		
Qy	1542	CAGCTTCATATATATTAATCTACTTCGAAAGATGATGATTTTATAGATGCGCTGTG	1601		
Db	80	CAGCTTCATATATATTAATCTACTTCGAAAGATGATGATTTTATAGATGCGCTGTG	21		
Qy	1602	AAAGACAAAAAATAAAAAA 1621			
Db	20	AAAAAATAAAAAAATAAAAAA 1			

RESULT 4	CK210372				
LOCUS	CK210372	1160 bp	mRNA	linear	EST 08-DEC-2003
DEFINITION	FGAS022177	Triticum aestivum FGAS: Library 5 GATE 7	Triticum		
ACCESSION	CK210372	aestivum cDNA, mRNA sequence.			
VERSION	CK210372.1	GI:39572762			
KEYWORDS	EST.				
SOURCE	Triticum aestivum (bread wheat)				
ORGANISM	Triticum aestivum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
AUTHORS	Allard, P., Crosby, W.L., Danyluk, J., Eudes, E., Frick, M., Gaudet, D., Genswein, B., Grat, R., Gulick, P., Hrycan, L.D., Laroche, A., Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, F., Penniket, C., Roach, J.L. and Sarhan, F.				
TITLE	Functional Genomics of Abiotic Stress In Wheat and Canola Crops				

JOURNAL COMMENT

Unpublished (2003)  
Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_estcs.usask.ca  
This sequence is the direct result of the Base calling software Phred (default parameters). It is the raw base calls. To aid in the identification of the high quality insert the software Lucy (default parameters) has been run on this sequence. Lucy identified the region [18,834].  
Plate: USB022 row: K column: 15.  
Location/Qualifiers  
1. 1160  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: Library 5 GATE 7"  
/notes="Vector: pCMV.SPORT6; Crown and developmental stages of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from lcn crown sections after 30 days of cold acclimation. The second is from lcn crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

FEATURES source

ORIGIN

Query Match	34.6%;	Score 562;	DB 7;	Length 1160;
Best Local Similarity	72.5%;	Pred. No. 4.4e-152;		
Matches 810;	Conservative	0;	Mismatches 271;	Indels 36;
Gaps	5;			
Qy	25	CGCGCGCGCGGCAAGCAATCATCATCTTCAATAGCTCATCAATCTATTCCGATGACG	84	
Db	14	CGNGACCCAGCGTCCGCGAACAACAACCAACCAATCTACCAACCTCTCCATGCGCA	73	
Qy	85	ACCGTCCCCAAGCCACAGCAGGTGGCGCATCTACGGGCAAGGTGCGACCCCGCGCATC	144	
Db	74	ACCGCGCCACGTCGACACAGCTCGCGCGGGGCGGCTGGAGCTCCCTCCCGCGCGC	133	
Qy	145	GGCGCTTGTGGAGTTTGCCCGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	204	
Db	134	GGGTTCCGGGCGCTCAAGGACTCCGCGCGCGCTCCGCTCTCTCTCTCTCTCTCTCTCT	193	
Qy	205	CGCACTCTCGCTAGAGGTGAAGGCATCTGGAAACACAGTTTGGAAATCTACTTTCAAGT	264	
Db	194	CGCGCGCGCGCTAGAGGTGAAGGCATCTGGAAACACAGTTTGGAAATCTACTTTCAAGT	253	
Qy	265	GCAACCTATGTTGAATCTCATGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	324	
Db	254	GCAACTTATGGAGAGTCTCATGGAGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	313	
Qy	325	AGAAATCCACTCACTGAGGCGAGACCTACAAAGTTGAAGTTCGATCGAAGCGCGCGGAC	384	
Db	314	AGAAATCCACTCACTGAGGCGAGACCTACAAAGTTGAAGTTCGATCGAAGCGCGGAC	343	
Qy	385	AGCAAGATAACCTCCCAAGAGAGGAGACTGATACATGCAAAATTTCTGTGAGGACACAT	444	
Db	344	AGCAAGATAACCTCCCAAGAGAGGAGACTGATACATGCAAAATTTCTGTGAGGACACAT	403	
Qy	445	GAAGGGGTGATCTACTGTAAGCGCCAAATTTCTGTTATTGTTCCCAACACAGATCAAAAT	504	
Db	404	GAAGGAATGACCACTGGGACACCAATTTCTGTTATTGTTCCCAACACAGATCAAAAT	463	

```
QY 505 AGTGATCAGCGTGAATAGCAATGTGTACCGACCTTCTCATGACAGCAACTTATGAC 564
    |||||
Db 464 GTGTATTACACTGAATAGCTAGGGGTACAGACCTTCCTCCATCGGATTAATCTATGAC 523
    |||||
QY 565 TTCAAGTACGGTGTATAGACTGTACAGGAGGTGGAGGTCTCTCGGCGAGAAACCGTT 624
    |||||
Db 524 CTCAATACGGTGTATAGACTGTGACAGGAGGTGGAGGTCTCTCGGCGAGAAACCGTT 583
    |||||
QY 625 GGAAGGCTGCTGCAGGGGCCCTCCCAAGAAATTTCTTAAGCTCAAAATGTGGATTAGAG 684
    |||||
Db 584 GGAAGGCTGCTGCAGGAGCTGTGCAAAAGAAATTTCTTAAGCTCAAAATGTGGGTTAGAG 643
    |||||
QY 685 ATCTCTGCTGTTGTTTCCAAAGTGCATCAGGTGTGCTCCCAAGAACCGGTGATTAT 744
    |||||
Db 644 ATCTAGACTGTTGTTTCCAAAGTGCATCAGGTGTGCTCCCAAGAACCGGTGATTAT 703
    |||||
QY 745 GGGTCTGTAACTTTGAAACAGATAGAGACCAATCGTTAGATGTCCTGATCCAGAGTAC 804
    |||||
Db 704 GAAAGCTCTTACCTGGATCAGATAGAGACCAATCGTAGATGTCCTGATCCAGAAAT 763
    |||||
QY 805 GCAGAGAAATGATAGACCAATCGACAGATACAGATTTCGAGGGGATTCGGTTCGTGA 864
    |||||
Db 764 GCACAGAAATGATGATGCAATTTGATAAAGTACGAAATTAATGGGAAATTCGATTGGTGG 823
    |||||
QY 865 GTGATCACATGCTGCTAGAAAGCTTCTCTCGGGCTCGGTTCTCTGTTTCGACACAG 924
    |||||
Db 824 GTGCTCACATGATGCTCCAGAAATGCTCTCGGGCTTGGTTCTCTGTTTCGACAAA 883
    |||||
QY 925 CTGAAATCCGAATGCTGCAAAAGCTATGCTTCTTATTCCTCGAGCAACCGGTTTCGAGATT 984
    |||||
Db 884 CTGTGAGCTCTACTGCGCAAGGCTATGCTTCTTCTTCTGCAAGCAAGGGGTTGAGATC 943
    |||||
QY 985 GCGAGCGGATTCGCGGGGACCGACTTGACAGGAAGTGAGCATATGATGAGTTTATATG 1044
    |||||
Db 944 CGTCAATGGAATTCGAGGTACTGACCCCTACTGGAGTTGAG-ATAACGATAAGTTCTATATG 1002
    |||||
QY 1045 GATAAGGCTGGAAGTGTACGAGACACGAGCTAATCGCTCGGGTGTGTCGAGGAGGATA 1104
    |||||
Db 1003 GATTAGGTTGAATGGT--GAACACGAACCATCGCTC-GGCGGGTTCAGGAGGGATA 1059
    |||||
QY 1105 TCGAATGTTGAGATTGTGCACCTTCAAAGTTGCTTTTA 1141
    |||||
Db 1060 TCAATGGGGAACCTATT--TCTTCAAAGTTGCTTTAA 1094
    |||||
```

```
RESULT 5
CB654761
LOCUS
DEFINITION
  OSJNEc07113.f OSJNEc Oryza sativa (japonica cultivar-group) cDNA
  clone OSJNEc07113 5', mRNA sequence.
ACCESSION
  CB654761
VERSION
  CB654761.1 GI:29658486
KEYWORDS
  EST.
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
  1 (bases 1 to 763)
REFERENCE
  Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
  Kudrta,D., Dean,R., Soderlund C., Wing,R. and Wang,G.
  Large-scale identification of ESTs involved in the interaction
  between rice and Magnaporthe grisea
  Unpublished (2003)
  Contact: Rod Wing
  Arizona Genomics Institute
  University of Arizona
  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
  85721-0088, USA
  Tel: 520 626 3967
  Fax: 520 621 9288
  Email: http://genome.arizona.edu
```

```
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: 1 column: 13
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
  1..763
  /organism="Oryza sativa (japonica cultivar-group)"
  /mol_type="mRNA"
  /cultivar="Nipponbare"
  /db_xref="taxon:39947"
  /clone="OSJNEc07113"
  /tissue_type="Leaf"
  /dev_stage="3 week"
  /lab_host="DH10B"
  /clone_lib="OSJNEc"
  /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
  XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN
Query Match      33.8%; Score 550.2; DB 6; Length 763;
Best Local Similarity 82.6%; Pred. No. 1.1e-148;
Matches 630; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 415 GATACATGCAAAATTTCTGTGAGGACACATGAAAGGGGTGACTACTGGAACGCAATTTCTT 474
    |||||
Db 1 GACACTTGC AAAATTTCTTTGAGGACACATGAAAGGAATGACCACTGGGACCAATTTCA 60
    |||||
QY 475 GTTATTTGCCAAAACACAGATCAATAGCGGTGATCAACCGTGAANAATAGCCAAATGTGTAC 534
    |||||
Db 61 GTTTTGTGTCGGAACACACAGATCAGAGAGGGGTGATTACAGTGAANAATGCTTAAGSCCTAC 120
    |||||
QY 535 GCACCTTCTCATGACAGACCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGA 594
    |||||
Db 121 AGACCTTCACATGACAGATGCAACTTATGACTTCAAAATACGGTGTGAGACAGTGCAGGA 180
    |||||
QY 595 GGTGGAGGTCTCTCGGGCAGAAAAACCGTTCGAAGGGTGGCTGACAGGGCCCTCCCCAG 654
    |||||
Db 181 GGTGGAAGATCATCAGCAAGAGAGACCATTGGAGGGTGGCTGACAGGAGCTCTTCCAAAG 240
    |||||
QY 655 AAAATTTCTTAAGCTCAAAATGTGGATTAGAGATTTGTCTGTTTGTTCCTTTTCAAAAGTGCATCAG 714
    |||||
Db 241 AAAATTTCTTAAGCTCAAAATCTGGAGTAGAGATCTTGGGGTGTGTTGTCCTCAAGGTGCATCAA 300
    |||||
QY 715 GTTGTGCTCCAGAGACCGGTTGATTATGGTCTGTAACTTTTGGAAACAGATAGAGAC 774
    |||||
Db 301 GTTGTACTACCAAGAGATGCGGCTTGTATGACACTGTAAACAATGGAAACAGATAGAAAGC 360
    |||||
QY 775 AACATCGTTAGATGTCTGTAGTCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGA 834
    |||||
Db 361 AACATTTGTAGATGTCTGTAGTCCAGATATGCAAGAAGATGATTGATGCAATTCGATAAA 420
    |||||
QY 835 GTACGAGTTCCAGGGGATTCGGTTCGGTGGAGTGTACATCGCTCGCTAGAAACGTTCTCT 894
    |||||
Db 421 GTACGAGTTAGAGGTGATTTCGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
    |||||
QY 895 CGCGGGCTCGGTTCTCTGTTGTTGCAAGCTCGAATCGAACTGGCAAAAGCTATGCTT 954
    |||||
Db 481 CGTGGGATTGGCTCTCTCTGTTATTTGACAAAACCTTGAGGCTGAAATTTGGCGGAAAGCTATGCTT 540
    |||||
QY 955 TCTATTCCTCGAGCAACCGGTTTCGAGATTTCGACGCGGATTCGCCGGGACCGACTTTGACA 1014
    |||||
Db 541 TCTCTTCTCGAAGCAAGGGGTTTGGATTCGGAGTGGATTTTGCAGGTACTGACTACTACT 600
    |||||
QY 1015 GGAAGTGAGCATATATGATGAGTTTTTATATGTAAGGCTGGAAGTGTGAGGACACGGACT 1074
    |||||
Db 601 GGAAGTGAGCATATATGATGAGTTCTATATGATGAGGCTGGAATGTGAGAACAGAACT 660
    |||||
QY 1075 AATCGCTCGGTTGTGTCAGGGAGGATATCGAATGTTGAGATTTGTCGACTTCAAAGTT 1134
    |||||
Db 661 AATCGTTCAGGCGGTGTGTCAGGGAGGATATCAAAATGTTGAAATTTATATATCTTACAGATGA 720
    |||||
QY 1135 GCTTTTAGCCGACACCACTTATCGGGGTGAAACAGAACACCG 1177
    |||||
```

```
FEATURES
source
  /organism="Oryza sativa (japonica cultivar-group)"
  /mol_type="mRNA"
  /cultivar="Nipponbare"
  /db_xref="taxon:39947"
  /clone="OSJNEc07113"
  /tissue_type="Leaf"
  /dev_stage="3 week"
  /lab_host="DH10B"
  /clone_lib="OSJNEc"
  /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
  XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
```

[illegible]







```

/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL102H06"
/tissue_type="adult vegetative tissue"
/plasmid="pCWSVSPORT 6"
/complement(1..1538)
/gene="At1g48850"

```

## ORIGIN

```

Query Match      30.7%; Score 499.8; DB 3; Length 1538;
Best Local Similarity 66.6%; Pred. No. 6.7e-134;
Matches 759; Conservative 0; Mismatches 377; Indels 3; Gaps 3;

QY 216 CCTAGAGGTGAAGGACATCTGGAACACAGTGTGGAACACTACTTTTCAGTTTGCAACCTATGG 275
DB 269 CTTCCAGATACAGCTACTGGAAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTTGG 328

QY 276 TGAATCTCATGGGGTGGTGTGTTGTTGTTATTCAGTGGTTGTCCACCTAGAAATCCACT 335
DB 329 AGAATCACATGAGGAGGAGTGTGTTGTTATTCATTTGATGTTGTCTCTCGTATTCCTACT 388

QY 336 CACTGAGGACAGCTCAAGTTGAACATCGATCGAAGACGCGCCGACAGAGCAGAGATAC 395
DB 389 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGAGGCTGTGTGACAGCAGGATCAC 448

QY 396 CTCACAAGGAAGGAGCTATACATGCAAAATTCCTGTCAGGACACATGAAGGGGTGAC 455
DB 449 AACTCTAGAAAGAGACTGATCTGCGCGATATCGTCTGAGTCACTGAAGGAGATGAC 508

QY 456 TACTGGAAGCCCAATTTCTTTATTTGTTCCCAACACAGATCAAAATAGGACGTGATCACCG 515
DB 509 GACAGGAACACCTATCCATGTTGTTGTTACCAACACAGATCAGAGGAGTGTGATTACAG 568

QY 516 TGAATAGCAATGTGTACCGACCTTCTCATGACAGCAACTTATGATTCAGTACGCG 575
DB 569 TGAATGTGCGTTGCTATAGACCATCGATGCTGATGCAACTTATGATGATGATGATG 628

QY 576 TGTAGAGCTGTACAGGAGGTGGAGGTCTCGGCGCAAAACCGTTGGAAGGTGCG 635
DB 629 TGTGAGTCAAGTGGAGGTGGAGAGATCTTCAGCTAGAGACCATTTGGGAGAGTTCG 688

QY 636 TGAGGGGCGCTCCCAAGAAATTTCTTAAGCTCAAAATGTGATTAGAGATCTTGTGCTT 695
DB 689 TGTGAGCTTTGGSCAAG-AAATTTGAAGCAATTTGACGAACTGAGATTTCTGCTTA 747

QY 696 TGTTCCTCAAGTGATCA-GGTTGTCTCCAGAGACGCGGTTGATTTATGGGTCGTAA 754
DB 748 TGTCTCGCAAGTCCACCATTTGTTGTTACTCCAGAGAAATTTGTAGACACGAGATTTTA 807

QY 755 CTTTGGACAGATAGAGACATCGTTAGATGCTGATCCAGAGTACGACAGAGAGA 814
DB 808 CACTCGAACAGATAGAAATAACATTTGTGATGCTGATGCTTAAATCCGAGTATGCGAAGA 867

QY 815 TGATAGACCAATCGACAGATACAGTTTCAGGGGATTCGCTGCTGAGTGATGATCAAT 874
DB 868 TGATAGCTCGGATTTGATGCTGTGACAGCAAAAGGAACTCTGTTGGTGTGTTGACT 927

QY 875 GGGTCGTAGAAACGTTCTCCGCGGCTCGGTTCTCTGTGTGTGACAAAGCTCGAATCCG 934
DB 928 GCAATGTTTCGGAATGCTCCACGTGGCTTGGTACACCGGTTTTCGATAAACTTGAAGCAG 987

QY 935 AACTGGCAAAAGCTATGCTTTCTATCTCTGCGAGCAACGGGTTCCAGATGTCACCGAT 994
DB 988 AACTGGCAAAAGCTTTGTTATGCTGCTACCTGCAACTAAGGGATTTTCAGTTTGGAAACGCT 1047

QY 995 TCTCCGGACCCGACTTTGACAGGAAGTGAGCATAAATGATGAGTTTATATGATGATAGGCTG 1054
DB 1048 TTGACAGGTACCTTTTGTGACTGTTGTTGACACAAATGATGAGTTCTATACCGATGAAATG 1107

QY 1055 GAAGTGTGAGGACACGGAATAATCGCTCGGGTGGTGTGTCAGGGAGGGATATCGAATGTTG 1114
DB 1108 GAAGAATACGTACCAAGAACCAACCGATCTGGTGGAAATTCAGGGAGGGATCTCAATGCTG 1167

```

```

QY 1115 AGATTGTGCACCTTCAAGTGTGCTTTTAAGCCGAC-ACCATCTATCGGGTGAACAGAAC 1173
DB 1168 AATAATAAATAGATAGTAGCTTCAAGCCCAACATCAACAAATTTGGAAGGAGAAAT 1227

QY 1174 ACCGTGTCAAGGGAGCGTTCAGAACGTTTGAGCTTCTAGCAAGAGGCGCCATGACCATGC 1233
DB 1228 ACGGTAAACAGAGACAAGTAGAACCAGAAATGATTGCGCGTGGTCTCATGATCTCTTGT 1287

QY 1234 GTGCGCCCTTCAGCTGTTCTGTGTGGAATTCATGGCGCGTGTGCTCTCATGACACG 1293
DB 1288 GTTGTTCACAGAGCTGTCCCAATGTTGGAAGCAATGTTGGTCTAGTTCTTCTTGTGATCAA 1347

QY 1294 CTGATGGCGCAGCTGGCTCAGTGCAGATGTTTCGGCTCAATPACTGCACTTCAAGAACC 1352
DB 1348 TTGATGGCGCATACGCAATGCCATTTGTTTCCATAAATCCAGAGTTGCGGAACC 1406

RESULT 10
CF303160
LOCUS
DEFINITION
ABF1--01-104.g1 ABF3-overexpressing transgenic rice lambda phage
clone ABF1--01-104, mRNA sequence.
CF303160
ACCESSION
VERSION
CF303160.1 GI:33674921
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 723)
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioeconomics and Bioinformatics, Myongji University
Yongin, Gyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES
Location/Qualifiers
1..723
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="ABF1--01-104"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="ABF3-overexpressing transgenic rice lambda
phage cDNA library (ABF1)"
/note="vector: pBluescript SK(+); Site1: EcoRI; Site2:
XhoI; Leaf was dried for 2hrs. cDNA was inserted into
lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end
with XhoI site. mRNA was prepared from ABA-responsive
element binding transcription factor 3 overexpression
line."
ORIGIN
Query Match      30.7%; Score 498.6; DB 7; Length 723;
Best Local Similarity 83.5%; Pred. No. 1.2e-133;
Matches 589; Conservative 0; Mismatches 114; Indels 2; Gaps 2;

QY 343 GCAGACCTACAGTTGAACTCGATCGAAGACGCGCCGACAGAGCAGAAATTAACCTTCA 402
DB 20 GCAGATATGCAAGTAGAATCTGACCGGAGACGCGCCGACAGAGCAGAAATTAACCCCA 79

QY 403 AGGAAGGAGACTGATACATGCAAAATTTCTGTGAGGGACACATGAAGGGGTGACTACTGA 462

```



```
Db 80 AGAAGGAGACTGACACTTGCAGAAATCTTTTCAGGACACATGAGGAATGACCACATGGG 139
Qy 463 ACGCCAAATCTTGTATTGTCCTCCAAACACAGATCAAATAGGACAGTATCACCCTGAAATA 522
Db 140 ACACCAATTCATGTTTGTCCGAAACACAGATCAGAGAGGGGTGATTACAGTGAATG 199
Qy 523 GCCAATGTGTACCGACCTTCTCATGTCAGACGCAACTTATGACTTCAAGTAGCGGTGTAGA 582
Db 200 GCTAAGGCTACAGACCTTCTCATGTCAGATGCAACTTATGACTTCAATATACGGTGTAGA 259
Qy 583 GCTGTACAGGAGGTGGGAGGTCTCTCGGAGAGAAACCGTTTGAAGGGTGCCTGACAGG 642
Db 260 GCAGTGCAGGAGGTGGAGATCATCAGACAGAGACCATTTGAAGGGTGGCTGCAGGA 319
Qy 643 GCCCTCCCAAGAAAATCTTAAAGCTCAAATGTGGAATTAGAGATCTTGTGCTTTTTC 702
Db 320 GCTCTTGCAGAAAGAAAATCTTAAAGCTCAAATCTGGAGTAGAGATCTTGGCGGTTTGTGTC 379
Qy 703 AAAGTGCATCAGTTGTCTCCAGAACACGGGTTGATTTATGGTCTGTAACTTTGGAA 762
Db 380 AAGGTGCATCAAGTTGTACTACAGAAAGATCCCGTTGATTATGACACTGTAAACATGGAA 439
Qy 763 CAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGACAGAGAGATGATAGAC 822
Db 440 CAGATAGAAAGCAATGTTAGATGTCCTGATCCAGAAATATGCAGAAAGATGATGAT 499
Qy 823 GCAATCGACAGATACGAGTTTCAGAGGGAATCGGTCGGTG-GAGTGAATCAATCGCTGC 881
Db 500 GCAATCGAATAAGTACGAGTTAGAGGTGATTCGATTTGGTGGGTGATGATGCAATGTC 559
Qy 882 TAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTTGTTTCGACAAAGCTCGAATCCGAATGCG 941
Db 560 AAGAAATGTTCTCGTGGGATTTGGCGCTCTGTATTTGACAAACTTGAAGGCTGAATGGC 619
Qy 942 AAAGCTATGCTTCTTCTTCGCGCAACCGGGTTTCGAGATTCGACGCGGATTCGCGG 1001
Db 620 GAAGCTATGCTTCTCTCTCTGC-AGCAAGGGTTTGAGATCGGATGCGATGTTGCAAG 678
Qy 1002 GACCGACTTGACAGGAAGTGACATAATGATGAGTTTTATATGA 1046
Db 679 TACTGACTACACTGGAAGTGAGCATAATGATGAGTTCTATATGA 723
```

```
RESULT 11
BJ268020
LOCUS
DEFINITION
aestivum cDNA clone whol14j05 5', mRNA sequence.
ACCESSION
BJ268020
VERSION
BJ268020.1 GI:20088164
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 710)
AUTHORS
Ogihara, Y. and Murai, K.
TITLE
Expressed genes in Triticum aestivum
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsinin@genes.nig.ac.jp.
Location/Qualifiers
1..710
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"

FEATURES
source
```

```
/db_xref="taxon:4565"
/clone="whol14j05"
/tissue_type="pistil at heading date"
/dev_stage="Peekes", scale 10.5"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_oh"

ORIGIN

Query Match 30.0%; Score 487.6; DB 4; Length 710;
Best local Similarity 80.4%; Pred. No. 2e-130;
Matches 571; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 614 GAAAAACCCGTTGGAAGGGTGCCTGACGGGCGCCTCCCAAGAAAAATCTTTAAGCTCAAAAT 673
Db 1 GGAAGAAACCTTGGAGGGTAGCTGCAGAGACTGTTGCAAGAAAAATCTTTAAGCTGAAT 60
Qy 674 GTGGATTAGAGATCTTTGTGCTGTTTTTCCAAAGTGCATCAGGTTGTGCTCCCAAGAACG 733
Db 61 GTGGAGTAGAGATTCTAGCATTTGTTTCCAAAGTGCATCAAGTGGTACTCTTCTCGAAGACG 120
Qy 734 CGGTTGATTTAGGTCGTAACTTTTGGAAACAGATAGAGACCAACATCGTTAGATGTCCTG 793
Db 121 CAGTTGATTTAGAACTCTTTACCTCGATCAGATAGAGACCAACATTTGTAGATGTCCTG 180
Qy 794 ATCCAGAGTAGCGAGAGAGATGATAGACCAATCCGACAGATACGAGTTTCGAGGGGATT 853
Db 181 ATCCAGATATGCACAGAAAGATGATGATGCAATTTGATAAGTACGAGTTAATGGGAATT 240
Qy 854 CGGTCGGTGGAGTGAATCACAATGCTGCTAGAAACGTTCTCGCGGGCTCGGTTCTCCTG 913
Db 241 CGATTGCTGGGGTGCATGCAATGTCAGAAAAATGTTCTCGTGGGGCTTGGCTCTCCTG 300
Qy 914 TGTTGCAAGACTCGAATCCGAATCGAACTGGCAAAAGCTATGCTTTCTATTTCTCGAGGCAACG 973
Db 301 TATTTGCAAAACTTGAAGCTCTACTGGCAAAAGGCTATGCTTTCTCTTCTCGCAAGCAAGG 360
Qy 974 GGTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTTGACAGGAAGTGAGCATATGATG 1033
Db 361 GGTTCGAGATCGGTAGTGGGATTTGAGTACTGACCTAACTGGAAGTGAGCATTAACGATG 420
Qy 1034 AGTTTATATGATGAAAGCTGGAAGTGTGAGGACACGGACTAAATCGCTCGGGTGGTGTGC 1093
Db 421 AGTTCTATATGGACGAGGCTGGAAATGTAAGAACACGAAACCAATCGCTCGGGCGGTGTAC 480
Qy 1094 AGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAGTTGCTTTTAAAGCGACACCAT 1153
Db 481 AGGAGGGATATCAAATGGTGAACCTATATACCTTCAAAGTAGCTTTCAAGCAACACGAA 540
Qy 1154 CTATCGGGGTGAAACACAAACACCGTGTCAAGGGAGCGTCAGAACGTTGAGCTTCTAGCAA 1213
Db 541 CTATGGGAAGAGCAAAATACCTGTAACAGGGATCATGAGGATATCGAACTTCTGACAA 600
Qy 1214 GAGGGCCCATGACCCATCGCTCGCCCTCGAGCTGTTCTGTGGTGGAAATCCAAATGCCCG 1273
Db 601 GGGGTGCCCATGACCCATGTTGTCCTCGGGCTGTCCAAATGCTGGAGACGATGCTG 660
Qy 1274 CGTTGTCCTCATGCGACAGCTGATGCGGACGCTGCGCTCAGTGCAGATG 1323
Db 661 CATTTGGTCTCATGGACCACTGATGGCACAATGTTGCTCAGTGGCAGATG 710

RESULT 12
BU099249
LOCUS
DEFINITION
WHE3305_C09_E17S Chinese Spring wheat drought stressed root cDNA
library Triticum aestivum cDNA clone WHE3305_C09_E17, mRNA
sequence.
ACCESSION
BU099249
VERSION
BU099249.1 GI:22546938
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

Poideae; Triticeae; Triticum.  
 1 (bases 1 to 702)  
 Anderson, O.D., Chao, S., Close, T.J., Crossman, C., Fenton, R.D.,  
 Lazo, G.R., Nguyen H.T., Pham, J., Rausch, C.J., Turuspekov, Y.,  
 Wilson, C., Woo, J., and Zhang, D.  
 The structure and function of the expressed portion of the wheat  
 genomes - Chinese Spring drought stressed root cDNA library  
 Unpublished (2002)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105959773  
 Fax: 5105959818  
 Email: oanderson@pw.usda.gov  
 Sequences have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: SK primer.

## FEATURES

## Source

Location/Qualifiers

1..702  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHB3305 C09 E17"  
 /tissue\_type="root"  
 /dev\_stage="Full tillering stage"  
 /lab\_host="E. coli SOUR"  
 /clone\_lib="Chinese Spring wheat drought stressed root  
 cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Plants were grown under  
 normal conditions, then drought stressed to 80%, 70% and  
 60% RWC at Texas Tech University (D. Zhang in HT Nguyen  
 lab). Total RNA was prepared separately for roots  
 collected at the three different drought conditions. Equal  
 amounts of total RNA were pooled from all three samples,  
 poly(A) RNA were purified, one cDNA library was made, and  
 the cDNA clones were in vivo excised to give pBluescript  
 SK(-) phagemids in the TJ Clouse lab at the University of  
 California, Riverside (Fenton, Turuspekov). Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

## ORIGIN

Query Match 29.5%; Score 480.4; DB 5; Length 702;  
 Best Local Similarity 80.5%; Pred. No. 2.5e-128;  
 Matches 562; Conservative 0; Mismatches 136; Indels 0; Gaps 0;  
 QY 634 GCTGAGGGGGCCCTCCCAAGAAAATCTTAAGCTCAAAATGGATTAGAGATCTTGTGG 693  
 DB 5 GTGAGTAGCTGTGGCAAGAAAATCTTAAGCTCAAAATGGATTAGAGATCTTGTGG 64  
 QY 694 TTGTTTCCAAAGTGCATCAGGTGTGCTCCAGAGACGCGGTGATTATGGGTCTGTA 753  
 DB 65 TTGTTTCCAAAGTGCATCAGGTGTGCTCCAGAGACGCGGTGATTATGGGTCTT 124  
 QY 754 ACTTTGGAACAGATAGAGAGCAACATCGTTAGATCTCTGATCCAGATGCCAGAGAG 813  
 DB 125 ACCCTGGATCAGATAGAGAGCAACATTTGTAGATGCTCTGATCCAGATATGCAGAG 184  
 QY 814 ATGATAGAGCAATGCAGAGTACGAGTTCGAGGGGATTCGGTGGTGGATGATCACA 873  
 DB 185 ATGATGATGCAATGATAAAGTACGAGTAAATGGGAATTCGATGGGTGGTGTGACA 244  
 QY 874 TCGCTCGCTAGAAAAGTTCCTCGCGGGCTCGGTCTCTCTGTGTTCCGACAAAGTCGAATCC 933  
 DB 245 TGCATTGCCAGAAATGTTCTCTCGTGGCTTGCTCTCTGTATTTGACAAACTGAAGCT 304  
 QY 934 GAACTGGCAAAAGCTATGCTTTCTATCTCTGCGAGCAACGGGTTTCGAGATTGGCAGCGGA 993  
 DB 305 CTACTGCGCAAGGCTATGCTTTCTCTCTCTGCAAGCAAGGGGTTTGAGATCGGTAGTGA 364

QY 994 TTGCGCGGACGCGACTTGACAGGAAGTGCAGATAATGATGAGTTTTATATGATAGGCT 1053  
 DB 365 TTTGAGGAGTACTGACCTAACTGGAAGTGCAGATAACGATGAGTTCTATATGACGAGCT 424  
 QY 1054 GGAAGTGTGACGACACGGAATAATCGCTCGGCTGGTGTGTCAGGGAGGAGATATCGAATGTT 1113  
 DB 425 GGAAGTGTGACGACACGGAATAATCGCTCGGCTGGTGTGTCAGGGAGGAGATATCAAATGCT 484  
 QY 1114 GAGATGTGCACTTCAAAAGTTGCTTTAAGCCGACACCATCTATCGGGGTGAAACAGAC 1173  
 DB 485 GAACTATATATCTTCAAAAGTACCTTTCAAGCCAAACAGCAACTATTTGGGAAGAGCAAAAT 544  
 QY 1174 ACCGTGTCAAGGGAGCGTCAGAACTGTCAGCTTTCTAGCAAGAGGCGGCATCACCATGC 1233  
 DB 545 ACTGTAAACAGGGATCATGAGGATATCGAATCTTCTGACAAAGGGTTCGCATGACCCATGT 604  
 QY 1234 GTGCGCCCTCGAGCTGTTCTGTGTGGAATCCATGCGCGGTGGTCTCTCATGAGACCAAG 1293  
 DB 605 GTTGTCCCTCGGCTGTTCCAATGGTGAGACGATGGCTGCATTGGTCTCTCATGAGACCAAG 664  
 QY 1294 CTGATGGCGCACGTCGCTCAGTGGAGATGTTCCGGCT 1331  
 DB 665 CTGATGGCACATGTTGCTCAGTGGAGATGTTCCGGCT 702

## RESULT 13

## CAL89902

LOCUS CAL89902 864 bp mRNA linear EST 24-SEP-2003  
 DEFINITION SCCLLRIC06B07.g LRI Saccharum officinarum cDNA clone SCCLLRIC06B07  
 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

CAL89902  
 CAL89902.1 GI:35133533  
 EST.  
 Saccharum officinarum  
 Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
 complex.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 864)  
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: C06 row: B column: 07

Seq primer: T7 Promoter Primer.

Location/Qualifiers

## FEATURES

## source

1..864  
 /organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCCLLRIC06B07"

/lab\_host="DH10B"

/clone\_lib="LRI"

/note="Organ: Leaf roll from field grown adult plants  
 (large insert library); Vector: pSport1; Site 1: SalI;  
 Site 2: NotI; An unidirectional cDNA library generated  
 from [leaf roll from field grown adult plants (large  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using Superscript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be

obtained at <http://sucest.lad.ic.unicamp.br/public>"

Query Match	29.4%	Score 477.4;	DB 6;	Length 864;
Best Local Similarity	82.5%;	Pred. No. 2e-127;		
Matches 547;	Conservative 0;	Mismatches 116;	Indels 0;	Gaps 0;
Qy	701	CCAAAGTGCATCAGGTTGTCTCCAGAGAAGCGGGTTGATTATGGCTCTCTAACTTTGG	760	
Db	1	CTAAAGTGCATCAAGTGTCTTCCAGAAGATGCAGTTGATTATGAGCATGTAACTTTGG	60	
Qy	761	AACAGATAGAGAGCAACATCTTTAGATGTCTCTGATCCAGAGTACGCAGAGAAGATGATAG	820	
Db	61	AACAGATAGAGAGCAACATCTGTAGATGTCTCTGATCCAGAATAATGCAGAGAAGATGATG	120	
Qy	821	ACGCAATTCGACAGATGACGATTCGAGGGGATTCGGTCGGTGGAGTGAATCACAATGCCTCG	880	
Db	121	CTGCCATTTGATAAAGTAGTACGAGTTAGAGGAGATTCAAATTTGGTGGGGTGGTACATATGC	180	
Qy	881	CTAGAAACGTTCTCCGGGGCTCGGTTCTCTCTGTGTTCCGACAGCTCGAATCCGAATCG	940	
Db	181	CAAGAAACGTCCTCCGTGGGTCTTGCGTCTCTCTGTGTTTGACAAACTTTGAAGCTGAAC	240	
Qy	941	CAAAAGCTATTCGTTTCTTATTCCTCGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCG	1000	
Db	241	CAAAAGCTATGCTTTCTCTTCCTCGAAGCAAGGGGTTTGAGATTGGCAGTGGGTTTGGTG	300	
Qy	1001	GGACCGACTTGACAGGAAGTGAGCATTAATGATGATGTTTATATGGAATAAGCTCGGAATG	1060	
Db	301	GTCAGAGCTTTACTGGAAGTGAGCATTAACGATGAGTTCTATATGATGAGGCTGGAATG	360	
Qy	1061	TCAGGACACGCACTAATCGCTCGGGTGGTGTGCAGGGAGGGATATCGAATGTTGAGATG	1120	
Db	361	TGCGGACACGAACTAATCGCTTCAGCGCGTGTTCAGGGAGGGATATCGAATGTTGAAATTA	420	
Qy	1121	TGCACTTCAAAAGTTGCTTTTAAAGCGCACACCATATCTATCGGGGTGAAAACAGAACCCGT	1180	
Db	421	TTTACTTCAAAAGTGGCTTTTAAAGCCACAGCAACTATCGGAAAGAAACAAATACTGTGT	480	
Qy	1181	CAAGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGCTCGCC	1240	
Db	481	CAAGGAGCATGAGGATGTTGAACTTTTGGCAAGGGGCGCCATGACCCGTGTGTTGCTCC	540	
Qy	1241	CTCAGAGCTGTTCCGTGTGTAATCCATGGCCGGTGGTTCCTCATGACCAGCTGATGG	1300	
Db	541	CTCAGAGCTGTTCCATGTGTGGAATCAATGGCTGGCTGGTCTCTGATGACCGAGCTCATGG	600	
Qy	1301	CGCAGTGGCTCAGTGGAGATGTTTCGGCTCAATAATGCACTTCAAGAAACAGTTCGGCT	1360	
Db	601	CGCATATTTGGTCAATGTGAGATGTTTCGGGTGAACCTTTCGGCTACAGAGAGCCCATGGT	660	
Qy	1361	CTT 1363		
Db	661	CTT 663		

RESULT 14	
CD870292	
LOCUS	694 bp mRNA linear EST 11-JUL-2003
DEFINITION	AZ02.113P12F001120 AZO2 Triticum aestivum cDNA clone AZO2113P12,
	mRNA sequence.
ACCESSION	CD870292
VERSION	CD870292.1 GI:32554108
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum Tritium aestivum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
	Poideae; Triticeae; Triticum. 1 (bases 1 to 694) Genopiante.
REFERENCE	Genoplante, a major partnership french program in plant genomics
AUTHORS	Unpublished (2003)
TITLE	Unpublished (2003)
JOURNAL	

COMMENT	CONTACT: GENOPLANTE Genoplatne 93, rue Henri Rochefort 91025 EVRY CEDEX France Tel: 33 1 69 47 54 00 Fax: 33 1 69 47 54 10 This sequence has been generated in the framework of the french plant genomics programme 'Genoplatne' (http://www.genoplatne.com and http://genoplatne-info.infobiogen.fr).
FEATURES	Location/Qualifiers 1..694 /organism="Triticum aestivum" /mol_type="mRNA" /cultivar="recital" /db_xref="taxon:4565" /clone="AZO2113P12" /tissue_type="root" /clone_lib="AZO2"
ORIGIN	Query Match 29.2%; Score 475.6; DB 6; Length 694; Best Local Similarity 81.3%; Pred. No. 6.2e-127; Matches 564; Conservative 0; Mismatches 129; Indels 1; Gaps 1;  QY 331 CCACCTACTGAGCGACCTACAAGTTGAATCGATCGAAGACGGCCGACAGAGCAGA 390 DB       2 CCACTCACCGAGGAAGACATGCAAGGAGACCTTTGATCGAAGGGCGGCAGTCCAGAGCAGA 61  QY 391 ATAACTCCACAGGAAGGAGACTGATACATGCAAAATCTGTGAGGACACACATGAAGGG 450 DB       62 ATAAACACCCAGGAAGGAGACCGATCTTTGTAATAATCTTTTCAGGACATATGAAGGA 121  QY 451 GTGACTACTGGAAAGCCAAATCTTTGTTATTTGTCCCAAACACAGATCAAAATAGGCAGTGT 510 DB       122 ATGACCACTGGGACGCCGATTCATGTTTTTGTCCAAACACGGATCAAGAGGGGGTGT 181  QY 511 CACGGTCAAAATAGCCAATGTGTACCGACCTTCTCATGCAGACGCCCACTTATGACTTCAAG 570 DB       182 TACACTGAAATGGCTAAGGGCGTACAGACCTTCCCATCGGATTTGACTTATGACCTCAAG 241  QY 571 TACGGTCTTAGAGCTGTGACAGGAGGTGGGAGGTCTCCGGCAGAAAAACCGTTTGAAGG 630 DB       242 TACGGTGTAGACTGTGTTCAGGAGGTGGAGGTCTCCGCAAGAGAAAACCATTTGAAGG 301  QY 631 GTGGCTCGAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGATATAGATCTTG 690 DB       302 GTAGCTCGAGAGCTGTGTGCAAGAAAAATCTTAAAGCTGAAATGTGAGTAGAGATTTCTA 361  QY 691 TCGTTTTGTTTCCAAAGTGCAATCAGTTGTGCTCCCAAGACCGCGTTGATTATGGGTCT 750 DB       362 GCATTTGTTTCCAAAGTGCAATCAGTTGGTACTTCTCTGAAGACCGAGTTGATTTGAAACT 421  QY 751 GTAACCTTTGAAACAGATAGAGAGCAACATCGTTAGATGTCTGTATCCAGAGTAGCCAGAG 810 DB       422 CTTACCTCGATCAGATAGAGAGCAATTTGTAGATGTCTGTATCCAGAAATATGCACAG 481  QY 811 AAGATGATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTCCGTGGAGTGATC 870 DB       482 AAGATGATTGATGCAATTGATTAAGTACGAGTTAAATGCGGAATTCGATTTGGTGGGGTGGTC 541  QY 871 ACATGGCTCGTAGAAACGTTCTCCGGGGCTGGTTCTCTGTGTTTCGACAAAGCTCGAA 930 DB       542 ACATGCAATTGCGAGAAATGTTCCTCGTGGGGCTTGGCTCTCTCTGTATTGACAACTTTGAA 601  QY 931 TCCGAACCTGGCAAAAGCTATGCTTTCTATTCTCTCGAGCAACGGGTTTCGAGATTGGCAGC 990 DB       602 GCTCTACTGGCAAAAGGCTATGCTTTCTTCTCTCTGC-AGCAAGGGGTTTGAGATCGGTAGT 660  QY 991 GGATTCGCCGGGACCGGACTTTGACAGGAAGTGAGC 1024 DB       661 GGATTTGCAGTACTGACCTTAACCTGGAAGTGAGC 694

RESULT 15

CAL124299  
 LOCUS SCQGLR1086A09.g 690 bp mRNA linear EST 23-SEP-2003  
 DEFINITION Saccharum officinarum cDNA clone SCQGLR1086A09  
 5' mRNA sequence.  
 ACCESSION CAL124299  
 VERSION CAL124299.1 GI:34977607  
 SOURCE EST.  
 ORGANISM Saccharum officinarum  
 Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
 complex.  
 1 (bases 1 to 690)  
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccenter.fcav.unesp.br  
 Plate: 086 row: A column: 09  
 Seq primer: 17 Promoter Primer.  
 Location/Qualifiers  
 1..690  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCQGLR1086A09"  
 /lab\_host="DH10B"  
 /clone\_lib="LRI1"  
 /note="Organ: Leaf roll from field grown adult plants  
 (large insert library); Vector: pSport1; Site 1: SalI;  
 Site 2: NotI; An unidirectional cDNA library generated  
 from [leaf roll from field grown adult plants (large  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be  
 obtained at <http://sucest.lad.ic.unicamp.br/public>"

## FEATURES

source

Search completed: August 27, 2005, 04:20:17

Job time : 4912 secs

## ORIGIN

Query Match 29.1%; Score 473; DB 6; Length 690;  
 Best Local Similarity 82.2%; Pred. No. 3.6e-126;  
 Matches 567; Conservative 0; Mismatches 121; Indels 2; Gaps 2;  
 QY 538 CCTTCTCATGACGACCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGT 597  
 DB 1 CCATCCCATGACAGTCAACCTATGACTTCAAGTATGGTGTACAGAGCTGTGACGGAGGC 60  
 QY 598 GGGAGGTCTCGGGCAGAAAAACCGTTGGAAGGGTGGCTGACAGGGGCCCTCCCAAGAAA 657  
 DB 61 GGAAGGTCTATCGCCAGAGAAACCATTTGGCAGGGTGGCTGACAGAGCTCTTGCAAGAAA 120  
 QY 658 ATTCTTAAGCTCAATGTGGATTAGAGATCTTGTGTTTGTTCCTCAAGTGCATCAGGTT 717  
 DB 121 ATTCGAAGCTCAATTCAGGAGTGAGATTTTGGCATTTGTTCCTCAAGTGCATCAGGTC 180  
 QY 718 GTGCTCCAGAGACCGGCTTATGGTCTGTAACTTTGGAACAGATAGAGCAAC 777  
 DB 181 GTACTTCCAGAGATGCAGTTGATTATGAGACTGTAACTTGGAAACAGATAGAGCAAC 240  
 QY 778 ATCGTTAGATGTCCTGATCCAGAGTACCGAGAGAGATGATAGACGCAATCGACAGATA 837  
 DB 241 ATTGTTAGATGTCCTGATCCAGAAATATGACAGAGAGATGATTGTGTCCTGCAATGATAAGTA 300

QY 838 CGAGTTTCAGAGGGATTTCGGTTCGGTGGAGCTGATCACATCGCTCGCTAGAAAAGTTCTCTCGC 897  
 DB 301 COAGTTAGAGAGATTCAATTGGTGGGTGCTGCATGATGATTCGAAGAAACGTCTCCTCGT 360  
 QY 898 GGGCTCGGTTCTTCCTGTGTTTCGACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTTTCT 957  
 DB 361 GGTCTTGGCTCTCCTGTTTTTGGACAAACTTGAAGCTGAACCTGGCAAAAGCTATGCTTTCT 420  
 QY 958 ATTCTCGAGCAACCGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTGAAGAGA 1017  
 DB 421 CTTCCTGCAAGCAAGGGGTTTGAGATTGGCAGTGGGTTTGTGGTACGGACTTTTACTGGA 480  
 QY 1018 AGTGAGCATATAATGATGAGTTTTTATATGATAAGGCTGGAAGTGTCCAGCACACGGACTAAT 1077  
 DB 481 AGTGAGCATATAAGATGAGTTCTATATGATGAGGCTGGAATGTCCGACACAGAACTAAA 540  
 QY 1078 CGCTCGGTTGCTGTCAGGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGC- 1136  
 DB 541 TCCTCACGCGGTGTTTCAGGGAGGGATATCGAATGTTGGAATTTATTTACTTCCAAAGTGGCT 600  
 QY 1137 TTTTAAGCCGACACCATCTATTCGGG-GTGAACACAGAACACCGTGTCAAGGGAGCGTCA 1195  
 DB 601 TTTTAAGCCAAACAGCAACTATCGGGAGAGAAACAAAATACTGTGTCAAGGGAGCATGANG 660  
 QY 1196 ACGTTGAGCTTCTAGCAAGAGGGCGCCATG 1225  
 DB 661 ATGTGGAATTTTGGCAAGGGGGGCCCATG 690

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:24:14 ; Search time 305 Seconds  
(without alignments)

Title: US-10-677-179-7

Perfect score: 1626

Sequence: 1 gcacgagctcagcttcgtct.....caaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY NUC

dapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 24055568

Minimum DB seq length: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*

```
1: /cgn2_6/ptodata/1/ina/5A COMB.seg:*
```

```
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
```

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*

```
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
```

```
5: /cgn2_6/ptodata/1/ina/PTCTUS COMB, see
6: /cgn2_6/ptodata/1/ina/bacif11.ea1
```

6: /cgnz\_6/prodata/1/1na/backstage1.be

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score			Query Match	Length	DB	ID	Description
1	1626	100.0	1626	4	US-09-743-207-7			Sequence 7, Appli
2	870.4	53.5	1635	4	US-09-743-207-1			Sequence 11, Appli
3	621	38.2	966	4	US-09-743-207-11			Sequence 5, Appli
4	500.6	30.8	1015	4	US-09-743-207-5			Sequence 1, Appli
5	274	16.9	1053	3	US-09-610-040-1			Sequence 7, Appli
6	274	16.9	1053	3	US-09-610-040-7			Sequence 7, Appli
7	274	16.9	1053	4	US-10-267-763-1			Sequence 7, Appli
8	274	16.9	1053	4	US-10-267-763-7			Sequence 7, Appli
9	203.8	12.5	1080	4	US-09-248-796A-4099			Sequence 4099, Ap
10	185	11.4	300	4	US-09-313-294A-4734			Sequence 4734, Ap
C 11	180.2	11.1	640681	4	US-09-790-988-1			Sequence 1, Appli
12	179.6	11.0	1089	4	US-09-543-681A-120			Sequence 120, Appli
13	178	10.9	1664976	4	US-08-916-41B-1			Sequence 1, Appli
14	178	10.9	1664976	4	US-09-692-570-1			Sequence 1, Appli
15	170	10.5	1239	3	US-09-064-693A-20			Sequence 20, Appli
16	170	10.5	4530	3	US-09-064-693A-26			Sequence 26, Appli
C 17	169.6	10.4	6464	3	US-09-221-017B-168			Sequence 168, Appli
18	169.4	10.4	541	4	US-09-743-207-13			Sequence 13, Appli
19	165	10.1	1104	4	US-09-328-352-3536			Sequence 3536, Appli
20	160	9.8	1233	4	US-09-252-991A-13477			Sequence 13477, A
C 21	160	9.8	1263	4	US-09-252-991A-13857			Sequence 13857, A
22	159.6	9.8	1134	4	US-09-489-039A-5375			Sequence 5375, Ap
23	157	9.7	1140	4	US-09-252-991A-13588			Sequence 13588, A
24	156.4	9.6	1830121	4	US-09-557-884-1			Sequence 1, Appli
25	156.4	9.6	1830121	4	US-09-643-990A-1			Sequence 9, Appli
26	139.8	8.6	479	4	US-09-743-207-9			Sequence 12, Appli
27	136.2	8.4	21410	4	US-09-596-002-12			Sequence 12, Appli

Db 301 TGTGTTATCAGTGGTGTGTCACCTAGAAATTCACCTACCTAGGCGAGACCTACAAAGTTGAA 360  
Qy 361 CTCGATCGAAGACGCCCGGACAGAGCAGAGATTAACCTCCACAGGAAGAGACTGATACAA 420  
Db 361 CTCGATCGAAGACGCCCGGACAGAGCAGAGATTAACCTCCACAGGAAGAGACTGATACAA 420  
Qy 421 TSCAAAATTCGTCTAGGAGACATCAAGGGGTGACTACTGGAAGCCCAATTCCTGTTATT 480  
Db 421 TSCAAAATTCGTCTAGGAGACATCAAGGGGTGACTACTGGAAGCCCAATTCCTGTTATT 480  
Qy 481 GTCCAAAACACAGATCAAAATAGCAGTATCAACCGTGAAATAGCCAAATGTGTACCGACCT 540  
Db 481 GTCCAAAACACAGATCAAAATAGCAGTATCAACCGTGAAATAGCCAAATGTGTACCGACCT 540  
Qy 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACCGTGTAGAGCTGTATACAGGAGGTGG 600  
Db 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACCGTGTAGAGCTGTATACAGGAGGTGG 600  
Qy 601 AGGTCCTCGGGCAGAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAATTT 660  
Db 601 AGGTCCTCGGGCAGAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAATTT 660  
Qy 661 CTTAAGCTCAAAATGTGGATTAGAGATCTTGTGTTTTCAAAAGTGCATCAGGTTGTG 720  
Db 661 CTTAAGCTCAAAATGTGGATTAGAGATCTTGTGTTTTCAAAAGTGCATCAGGTTGTG 720  
Qy 721 CTCCAGAGACGCCGTTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780  
Db 721 CTCCAGAGACGCCGTTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780  
Qy 781 GTTACATGTCCTGATCCAGAGTACCCAGAGAGATAGAGCGCAATCGACAGAGTACGA 840  
Db 781 GTTACATGTCCTGATCCAGAGTACCCAGAGAGATAGAGCGCAATCGACAGAGTACGA 840  
Qy 841 GTTCAGGGGATTCGGTCCGTGGAAGTATCATCGTCCGTAGAAAAGTTCTCCGCGGG 900  
Db 841 GTTCAGGGGATTCGGTCCGTGGAAGTATCATCGTCCGTAGAAAAGTTCTCCGCGGG 900  
Qy 901 CTCGGTCTTCGTGTTTGCACAGCTCGAATCCGAACTGGCAAAAGCTATGCTTCTATT 960  
Db 901 CTCGGTCTTCGTGTTTGCACAGCTCGAATCCGAACTGGCAAAAGCTATGCTTCTATT 960  
Qy 961 CTTGCGAGCAACGGGTTGCAGATTCGCGAGCGGATTCGCGGACCGACTTCACAGGAAGT 1020  
Db 961 CTTGCGAGCAACGGGTTGCAGATTCGCGAGCGGATTCGCGGACCGACTTCACAGGAAGT 1020  
Qy 1021 GAGCATAATGATGAGTATATGGAATGAGGCTGGAAGTGTGAGGACGCACTAATCGC 1080  
Db 1021 GAGCATAATGATGAGTATATGGAATGAGGCTGGAAGTGTGAGGACGCACTAATCGC 1080  
Qy 1081 TCGGCTGTGTGACGGAGGGATTCGAAATGTTGAGATTTGCACTTCAAAGTTGCTTTT 1140  
Db 1081 TCGGCTGTGTGACGGAGGGATTCGAAATGTTGAGATTTGCACTTCAAAGTTGCTTTT 1140  
Qy 1141 AAGCGACACCACTTATCGGGGTGAAACAGAACCCGTGTCAAGGGAGCGTCAGAACGTT 1200  
Db 1141 AAGCGACACCACTTATCGGGGTGAAACAGAACCCGTGTCAAGGGAGCGTCAGAACGTT 1200  
Qy 1201 GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTCCGCCCTCGAGCTGTTCTGTGTG 1260  
Db 1201 GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTCCGCCCTCGAGCTGTTCTGTGTG 1260  
Qy 1261 GAATCCATGCGCGGTTGCTCTCATGGAACAGCTGATGGCGCAGCTGGCTCAGTGCAG 1320  
Db 1261 GAATCCATGCGCGGTTGCTCTCATGGAACAGCTGATGGCGCAGCTGGCTCAGTGCAG 1320  
Qy 1321 ATGTTGCGCTCAATCTACTGCACTTCAAGAACCCAGTTGGCTCTTTTACAGAGGCGAGC 1380  
Db 1321 ATGTTGCGCTCAATCTACTGCACTTCAAGAACCCAGTTGGCTCTTTTACAGAGGCGAGC 1380  
Qy 1381 ACACCTGATGAGCTCGCGCAAAATTTTATCATATAGTAAATAGTAGCTCAAGCGT 1440  
Db 1381 ACACCTGATGAGCTCGCGCAAAATTTTATCATATAGTAAATAGTAGCTCAAGCGT 1440

Qy 1441 GGCTTGGTTTCTGTCTCTTGACCGTAGTCTTTGTTTTTTTTTTTTTCCCGCAAGTGTGATG 1500  
Db 1441 GGCTTGGTTTCTGTCTCTTGACCGTAGTCTTTGTTTTTTTTTTTTTCCCGCAAGTGTGATG 1500  
Qy 1501 CGATGAAGTGAATAAGGCACCTTGGTTTCTGTGCAATTTGTACAGCTTTTCATATAATGTAA 1560  
Db 1501 CGATGAAGTGAATAAGGCACCTTGGTTTCTGTGCAATTTGTACAGCTTTTCATATAATGTAA 1560  
Qy 1561 TCTACTTCGAAGATGATGCAATTTTATAGATGGCTTGTGAAAGACAAAAA 1620  
Db 1561 TCTACTTCGAAGATGATGCAATTTTATAGATGGCTTGTGAAAGACAAAAA 1620  
Qy 1621 AAAAAA 1626  
Db 1621 AAAAAA 1626

## RESULT 2

US-09-743-207-1

; Sequence 1, Application US/09743207  
; Patent No. 6653531

; GENERAL INFORMATION:

; APPLICANT: Caboon, Rebecca E.

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Pember, Stephen O. Biosynthesis Enzymes

; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes

; FILE REFERENCE: BB-1159-A

; CURRENT APPLICATION NUMBER: US/09/743,207

; CURRENT FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: PCT/US99/16353

; PRIOR FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: 60/093,611

; PRIOR FILING DATE: 1998-07-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1635

; TYPE: DNA

; ORGANISM: Zea mays

US-09-743-207-1

Query Match 53.5%; Score 870.4; DB 4; Length 1635;

Best Local Similarity 80.5%; Pred.No. 1.1e-271;  
Matches 1031; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy 83 CGACCGTGCCCAAGCCACAGCAGGTGGCGCACTCACGGGCA CGGCTCGCACCCCGCGGA 142  
Db 59 CGCCCGTGTGCGAGCCCGCGGTGTCCGCCAGGGCGTCCACACGCTTCTCCCGCGGGA 118  
Qy 143 TCGGCGCTTGTGTGGAGTTTGGCCCGCAGCTCTCTCTCCCTCCGCTTCCGCTGACCGCT 202  
Db 119 TAGCGCGCTCCCGGAGTCCGCC--CCACGTCCTCCGCTTATCCGTCGGCGCGCTC 175  
Qy 203 GCAGCAGTGTGCTGCTAGAGTGAAGGCATCTGGAACACGTTTGGAACTACTTTTCAGG 262  
Db 176 GCGCGGCTTCAAGCTTAGAGTGAAGGCATCAGGAATGTTTCGGGAATCTACTTCAGG 235  
Qy 263 TTGCAACTATGTGAAATCTCATGGGGGTGGTGTGGTTGTGTTATCAGTGGTTGTCCAC 322  
Db 236 TTGCAACTATGCGAATCCATGAGGGGGGTGTGGTTGTGCTTATCAGTGGCTGCCAC 295  
Qy 323 CTAGAAATTCACCTCAGTGAAGGCAGACCTACAAGTTGAATCGATCGAAGACGGCCCGAC 382  
Db 296 CCAGAAATTCCTCTCAGTGAAGGCAGACATGCAAGTAGAACTCGATAGAAAGCGTCCGGTC 355  
Qy 383 AGACAGAAATAACCTCCACAAAGGAAGAGACTGATACATGCAAAATTCCTGTCAGGAGCAC 442  
Db 356 AAAGTAGAATTAACACCCCAAGAAAGAGACTGATACATGCAAAATTCATTCAGGAGCAC 415  
Qy 443 ATGAAGGGGTGACTACTGGAACCGCAATTCCTGTTATGTCCTCCCAACACAGATCAAAATAG 502  
Db 416 ATGATGGGATGACTACTGTTACACCAATTCACGCTCTTTTGTGTCCTCCCAACACAGATCAAAAGG 475



```

QY 503 GCAGTGATCAACGTTGAAATAGCAATGTGTACCGACCTTCTCATGCGACGCAACTTATG 562
Db      |||||
QY 476 GTGGTATGATACAGTGAATGTCTTAGCGGTACAGACCATCCATGATGACGATGCACTATG 535
Db      |||||
QY 563 ACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGAGGTCTCTGGGCGAGAAACCG 622
Db      |||||
QY 536 ACTTCAAGTATGAGGTGAGAGTGTGACGAGGAGGTGGAAGGTCTATGAGGAGGAGGAGG 595
Db      |||||
QY 623 TTGGAGGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 682
Db      |||||
QY 596 TTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 655
Db      |||||
QY 683 AGATCTTGTGCTTGTCTTCAAGTGTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 742
Db      |||||
QY 656 AGATCTTGTGCTTGTCTTCAAGTGTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 715
Db      |||||
QY 743 ATGGGTGTGTAATCTTTGGAAAGATAGAGAGCAACATCTTATGATGTCTCTGATCCAGAGT 802
Db      |||||
QY 716 ATGAGACTGTAACTTTGGAACATATAGAGAGCAACATCTTATGATGTCTCTGATCCAGAGT 775
Db      |||||
QY 803 ACGCAGAGAGGATGATAGAGCAACATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 862
Db      |||||
QY 776 ATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 835
Db      |||||
QY 863 GAGTGATCACATGCGGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 922
Db      |||||
QY 836 GGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 895
Db      |||||
QY 923 AGCTCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCG 982
Db      |||||
QY 896 AACTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 955
Db      |||||
QY 983 TTGGCAGCGGATTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1042
Db      |||||
QY 956 TTGGCAGTGGGTTCGCTGTGACGAGCTTTTACTGGAAGTGAAGTGAAGTGAAGTGAAGT 1015
Db      |||||
QY 1043 TGATTAAGCTGGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1102
Db      |||||
QY 1016 TGATTAAGCTGGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1075
Db      |||||
QY 1103 TATCGAATGTGAGATGTGATCTTCAAGTGTGCTTTTAAAGCGGAGGAGGAGGAGGAGGAG 1162
Db      |||||
QY 1076 TATCGAATGTGAGATGTGATCTTCAAGTGTGCTTTTAAAGCGGAGGAGGAGGAGGAGGAG 1135
Db      |||||
QY 1163 TGAACAGAGCAACCGGTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1222
Db      |||||
QY 1136 AGAAGCAAAATATCTGTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1195
Db      |||||
QY 1223 ATdACCCATGCTGCGCCCTCGAGCTGTTCCTGTGGAATCCATGCGCGGTGTGCTCC 1282
Db      |||||
QY 1196 ATdACCCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255
Db      |||||
QY 1283 TCATGAGCAGCTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1342
Db      |||||
QY 1256 TGATGAGCAGCTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1315
Db      |||||
QY 1343 TTCAAGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1362
Db      |||||
QY 1316 TACAAGAGGCCATTTGGCTCT 1335
Db      |||||

```

```

RESULT 3
US-09-743-207-11
; Sequence 11, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207

```

```

; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-743-207-11

```

```

Query Match 38.2%; Score 621; DB 4; Length 966;
Best Local Similarity 80.4%; Pred. No. 9e-191;
Matches 742; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

QY 75 TCCGATGACGACCGTGGCCCAAGCCACAGCAGGTGGCGCACTCAGCGGCGACGCTCGCAC 134
Db      |||||
QY 41 TCCGATGCGCGCGCAACGTCGTGCGAGCGGTGGCGCGCTCTCCCGCGCGCGCGCG 100
Db      |||||
QY 135 CGCGCGATCGCGCGCTTGTCTGGAGTTTGGCCCGAGCTCTCTCTCCCTCCGCTTCCGCT 194
Db      |||||
QY 101 CGCGCGGTTCGCGCGCTTCCCGAGTCCGCGCGCT-----TCCCTCCGCTTCTCCGT 154
Db      |||||
QY 195 GCACGCTGCGCGCACTGCTCGCTTAGAGGTCAAGGCACTCTGGAACAACAGTTTGGAAACTA 254
Db      |||||
QY 155 CGCGCGCGCGCGCGCTCTGCTAGAGTGAAGCGCTCTGCAAAATGATTTTGGAACTA 214
Db      |||||
QY 255 CTTTCAGGTTCGAACCTATGTTGAATCTCATGGGGGTGGTGTGGTGTGTTATCAGTGG 314
Db      |||||
QY 215 CTTTCAGGTTCGAACCTATGTTGAAGTCTCATGGAGGCGGTGTTGGTGTGTAATCAGTGG 274
Db      |||||
QY 315 TTGTCCACCTAGAAATTCACCTCAGTGGCGAGACCTCAAGTTGAACTCTGATCGAAGACG 374
Db      |||||
QY 275 ATGCCACCCAGAAATCCCACTTACTGAGAGCAGATATGCAAGTAGAACTCGACCCGAGACG 334
Db      |||||
QY 375 GCCCGGACAGACAGAGATAAATCTCCACAAGGAGGAGACTGATCATGTCGCAAAATTTCTGTC 434
Db      |||||
QY 335 GCGAGGCGAGAGAGAGATAAATCCACCCAGAGAGAGACTGACACTTGTGCAAAATTTCTTTC 394
Db      |||||
QY 435 AGGAGACATGAAGGGGTGACTACTTGGAAAGCCCAATTTCTTGTATTTGTCGCAAAACACAGA 494
Db      |||||
QY 395 AGGAGACATGAAGGAGTACCACTTGGGAGACCAATTCATGTTTGTGTCGCAAAACACAGA 454
Db      |||||
QY 495 TCAAAATAGGCAAGTATCACTGCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCAAGCGC 554
Db      |||||
QY 455 TCAGAGAGGGGTGATTTACAGTGAATGGCTTAAGGCGCTTACAGACCTTCCATATGCAAGTGC 514
Db      |||||
QY 555 AACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGGCAG 614
Db      |||||
QY 515 AACTTATGACTTCAAAATACGGTGTAGAGCTGTAGAGGAGTGGAGAGTGTGTAAGTGCATC 574
Db      |||||
QY 615 AAAAACCGTGTGGAAGGGTGGCTGCGAGGGGCGCTCCCAAGAGAAATTTCTTAAAGCTCAAAATG 674
Db      |||||
QY 575 AGAGACCATTTGGAAGGGTGGCTGCGAGGAGCTTTGCAAGAGAAATTTCTTAAAGCTCAAAATC 634
Db      |||||
QY 675 TGGATTAGAGATCTTGTGCTTTTGTTCGAAAGTGCATCAGGTTGTGCTCCAGAGAGCGC 734
Db      |||||
QY 635 TGGAGTAGAGATCTTGGCGTTTGTGTCGAGGTCATCAAGTTGTACTTACAGAGAGATGC 694
Db      |||||
QY 735 GGTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATCGTTAGATGTCCTGA 794
Db      |||||
QY 695 CGTTGATTATGACTGTAACTTTGGAACAGATAGAGCAACATCGTTAGATGTCCTGA 754
Db      |||||
QY 795 TCCAGAGTACGACAGAGAGATGATAGCGCAATCGACAGAGTACGAGTTCCAGGGGATTC 854
Db      |||||
QY 755 TCCAGAAATATGACAGAGAGATGATTTGATGCACTCGATTAAGTACGAGTTAGAGGTGATTC 814
Db      |||||
QY 855 GGTGCGGAGTATCATGCGTGTGCTGTAGAAAGCTTCTCGCGGGGTTCGGTTCTCTGT 914
Db      |||||
QY 815 GATTGGTGGTGTGCTGCTCATGCAATTTGCAAGAAATGTTCTCTCGTGGGATGGCTCTCTCTGT 874
Db      |||||

```

Qy 915 GTTCACAAGCTCGAATCCGAACGCGCAAAAGCTATGCTTTCTATTCTCTCGAGCAACGG 974  
Db |||||  
Qy 875 ATTTCACAACACTTGAGGCTGAATGGCGAAAGCTAAGCTTTCTCTTCTCGCAACGAGGG 934  
Db |||||  
Qy 975 GTTCGAGATTGGCAGCGGATTTCG 997  
Db |||||  
Qy 935 GTTTCGAGATCGGCACTGGGATTTCG 957  
Db |||||

## RESULT 4

US-09-743-207-5  
; Sequence 5, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 1015  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-743-207-5

Query Match 30.8%; Score 500.6; DB 4; Length 1015;  
Best Local Similarity 80.3%; Pred. No. 1.3e-151;  
Matches 587; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 631 GTGGCTGCAGGGCCCTCCCAAGAAATTTCTTAAGCTCAATGTGGATTAGAGATCTTG 690  
Db 5 GAGGCTGCAGGAGCTGTGTCAAGAAATTTCTTAAGCTGAATGTGGAGTAGAGATTTCTA 64  
Qy 691 TCGTTTGTTCAAAAGTCATCAGGTGTGCTCCCAAGAGACGCGGTTGATTATGGGTCT 750  
Db 65 GCATTTGTTCCAAAGTCATCAAGTGTGCTTCTGAAGACGCAAGTTGATTATGAACT 124  
Qy 751 GTAACCTTTGGAAACAGATAGAGCAACATCGTTAGATGCTCTGATCCAGAGTACGACAG 810  
Db 125 CTACCTCGATCAGATAGAGCAACATTTGTAGATGCTCTGATCCAGAAATATGCACAG 184  
Qy 811 AAGATGATAGACCAATCGACAGATGACAGTTCAGGGGATTCGGTGGGTGATGATC 870  
Db 185 AAGATGATGATGCAATTTGATAAGTACGAGTTAATGGGAATTCGATTGGGGTGGTC 244  
Qy 871 ACATCGCTCGCTAGAAACGTTCTCTCGGGGCTCGGTTCTCTGTGTTCCGACAAAGCTCGAA 930  
Db 245 ACATCGATTGCGAAGAAATGTTCTCTGGGCTTGCTCTCTGTATTTGACAAACTTGA 304  
Qy 931 TCCGAACCTGGCAAAAGCTATGCTTTCTATTCCTCGACGCAACGGGTTTCGAGATTGGCAGC 990  
Db 305 GCTCTACTGCGCAAGGCTATGCTTTCTCTCTCGCAAGCAAGGGGTTTTCGAGATCGGTAGT 364  
Qy 991 GGATTCGCGGACCGACTTGACAGAGTGAAGTATGATGATGATTTTATATGATGAAG 1050  
Db 365 GGATTTGCGAGTACTGACCTAATCGAAGTGAAGTATGATGATGATTTTATATGACGAG 424  
Qy 1051 GCTGGAAGTGTGAGACACGGAATTAATCGCTCGGGTGTGTGTCAGGAGGGATATCGAAT 1110  
Db 425 GCTGGAATGTGAAGAACACGACCAATCGCTCGGGGCTGTACAGGAGGGATATCAAT 484  
Qy 1111 GTTGAGATTGTCACTTCAAGTGTCTTTTAAGCCGACACCATCTATCGGGGTGAACAG 1170  
Db 485 GGTGAAACTATATATCTTCAAGTAGCTTTTCAAGCCCAACAGCAACTATTTGGGAAGAGCA 544

Qy 1171 AACACCGTGTCAAGGAGCGCTCAGAACGTTGAGCTTCTAGCAAGAGGCGCCCATGACCCA 1230  
Db |||||  
Qy 545 AATACTGTAAACAGGGATCATGAGGATATCGAACTTCTGAACAGGGTCCCATGACCCA 604  
Db |||||  
Qy 1231 TGGCTCGCCCTCGAGCTGTTCTCTGTGGTGAATTCATGGCCGCTGTGCTCTCATGGAC 1290  
Db |||||  
Qy 605 TGTGTCGCTCCTCGGGCTGTTCCAATGTTGGAGACGATGCTGCAATTGGTCTCATGGAC 664  
Db |||||  
Qy 1291 CAGCTGATGGCGACGCTGGCTCAGTGGAGATGTTCCGCTCAATACTGCACTTCAAGAA 1350  
Db |||||  
Qy 665 CAGCTGATGGCACATGTTGCTCAGTGGAGATGTTCCGCTGAACCTCGCCCTACAAGAA 724  
Db |||||  
Qy 1351 CCAGTTGGCTC 1361  
Db 725 CCAATCGGCTC 735

## RESULT 5

US-09-610-040-1/c  
; Sequence 1, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 1053  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1022)..(1023)  
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).  
; NAME/KEY: misc feature  
; LOCATION: (1025)..(1025)  
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).  
; NAME/KEY: misc feature  
; LOCATION: (1033)..(1034)  
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).  
; NAME/KEY: misc feature  
; LOCATION: (1044)..(1044)  
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).  
US-09-610-040-1

Query Match 16.9%; Score 274; DB 3; Length 1053;  
Best Local Similarity 64.0%; Pred. No. 6.1e-78;  
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGGTGAGGCATCTCGNAACACGTTTGGAACTACTTTCAGGTTCGAACCTATGG 275  
Db 748 CTTCCAGATCAAGCTACTGGAAGTTCAATATGGACTCATTTTCGAGTTTCAACTTTGG 689  
Qy 276 TGAATCTCATGGGGTGGTGTGGTTGTGTTATTCAGTGGTTGTCCACTAGAAATCCACT 335  
Db 688 AGAATCATATGGAGGAGGAGTTGGTTGTATCATTTGATGTTGTCTCTCGTATTCACCT 629  
Qy 336 CACTGAGGCGACCTCAAGTTTGAATCTCGATCGAAGACGGCCCGGACAGAGCAAGATAAC 395  
Db 628 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGAGGCGCTGTGTGAGAGCATCAC 569

```

; LOCATION: (31)..(32)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
US-09-610-040-7

Query Match      15.9%; Score 274; DB 3; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6.1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1

Qy 216 CCTAGAGGTGAAGGCATCTGGAACACAGTTTTGAAACACTCTTTTCAGGTTGCAACCTATGG 275
Db |||||
Qy 276 TGAATCTATGSGGGTGGTGTGGTGTGTATCAGTGGTGTGTCCACTAGAAATTCCTACT 335
Db |||||
Qy 366 AGAATCATCGGAGGAGGTGGTGTGTATCATGTAGTGGTGTCTCTCTCGTATTTCCTACT 425
Db |||||
Qy 336 CACTGAGGCAGACCTCAAGTTGAATCTCGATCGAAGACGGCCCGACAGACGACGAATAAC 395
Db |||||
Qy 426 TACTGAATCTGAATTTGGCAATTCGATCTCGATAGAAGGAGGCTCGTTCAGACGAGGATCAC 485
Db |||||
Qy 396 CTCACAAGGAAGGAGACTGATACATGCAAAATTTCTGTCAAGGACACATGAAAGGGGTGAC 455
Db |||||
Qy 486 AACTCTCTAGAAAAGAGACTGATACCTTCGCGGATATCGTCTGAGACTCTCTGGAAGAAATGCAC 545
Db |||||
Qy 456 TACTGGAGCGCCAAATCTCTGTATTGTATGTCCTCCAAACACAGATCAAAATGAGCACTGATCACCG 515
Db |||||
Qy 546 GACAGGAACACCTATTCATGTGTTGTGACCAACACAGATCAGAGAGGACTTGATTAACAG 605
Db |||||
Qy 516 TGAATAGCCAAATGTGTACCGACCTTCTCATGACAGACGCAACTTATGACTTCAAGTACGG 575
Db |||||
Qy 606 TGAATGTGCGTTGCGCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAGTATGG 665
Db |||||
Qy 576 TGTTAGAGCTGTACAGGAGAGTGGGAGGTCTCTGGGCGAGAAAAACCGTTGGAAGGGTGGC 635
Db |||||
Qy 666 TGTGAGATCAGTGCAGGCTGGAGGAAGATCTTCAGCTAGACAGACCAATTCGAAGAGTTGC 725
Db |||||
Qy 636 TGCAAGGGCCCTCCCAAGAAAAATCTTTAAGCTCAATGTGGATTAGAGATCTTGTGCGTT 695
Db |||||
Qy 726 TCTTGGAGCTTTGGGCAAGAGAAATTTTGAAGCAATTTGCAAGAACTCAGAAATCTTGCCTA 785
Db |||||
Qy 696 TGTTTCCAAAGTGCATCAGGTGTGCTCCCAAGAGACGCGGTTGATTAATGGGTCTGTAAC 755
Db |||||
Qy 786 TGTCTGCAAGTTTCAACAAATGTGTACTTTCAGAAGAAATTTGTCAGACGAGAAATTCACAC 845
Db |||||
Qy 756 TTTTGGACAGATAGAGAGCAACATCGTTAGATGTCCTGTATCCAGAGTACGCGAGAGAAAT 815
Db |||||
Qy 846 TCCG--ACAGATGAAAAATAACATTTGTTCAGAGCCCTAAATCCGAAATAATCCGAAAAAGA 903
Db |||||
Qy 816 GATTAGCGCAATCGACAGAGTACAGTTTCAGGGGGAATTCGGTCCGTGGAGTGATCACATG 875
Db |||||
Qy 904 GATTAATGCGGAATGAATCCGTCAAGACGATAGGAATCTTGTGGTGGGTTTTGACCTG 963
Db |||||
Qy 876 CGTCGCTAGAAA 887
Db |||||
Qy 964 CATTTGCGGAAA 975

```

RESULT 6  
US-09-610-040-7  
; Sequence 7, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Gurlach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS

```

, FILE REFERENCE: 9128.14
, CURRENT APPLICATION NUMBER: US/09/610,040
, CURRENT FILING DATE: 2000-07-05
, NUMBER OF SEQ ID NOS: 11
, SOFTWARE: PatentIn version 3.0
, SEQ ID NO 7
, LENGTH: 1053
, TYPE: DNA
, ORGANISM: Arabidopsis thaliana
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (10)..(10)
, OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
, NAME/KEY: misc feature
, LOCATION: (20)..(21)
, OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
, NAME/KEY: misc feature
, LOCATION: (29)..(29)
, OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
, NAME/KEY: misc feature

```

```
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1053)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-1

Query Match      16.9%; Score 274; DB 4; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6,1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGGTGAAGGCATCTGGAACACAGTTTGGAACTACTTTCAGGTTCGAACCTATGG 275
Db 748 CTTCCAGATACAAGCTACTGGAAGTTTCATATGGGACTCATTTTCGAGTTTCAACTTTGG 689

Qy 276 TGAATCTCATGGGGTGTGGTGTGGTATCATAGTGTGGTGTCCACCTAGAAATCCACT 335
Db 698 AGAATCACAATGGAGAGGAGTTGGTTGTATCATTTGATGGTTGTCTCTCGTATTTCCACT 629

Qy 336 CACTGAGGCAGACCTACAAAGTTGAATCGATCGATGAGAGCGCCGCGACAGACAGAAATAAC 395
Db 628 TACTGAAATCTGATTTTGCAATTCGATCTCGATAGAGAGGCTGTGTGACAGCAGATCAC 569

Qy 396 CTCCACAAGGAAGGAGACTGATACATGCAAAATTCCTGTGAGGACACATGATGAAGGGGTGAC 455
Db 568 AACTCCTAGAAAAGAGACTGATACCTTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGAC 509

Qy 456 TACTGGAAGCCAAATCTGTTATTTGTGCCAACAACAGATCAATCAATAGGAGTATCACCG 515
Db 508 GACAGGAACACCTATCCATGTGTTTGTATCCAAACAACAGATCAGAGAGGACTTGATACAG 449

Qy 516 TGAATAGCCAAATGTATCCGACCTTCTCATGACAGCGCAACTTATGACTTCAAGTACGG 575
Db 448 TGAATGTGCGTTGCTATAGACCAATCGCATGCTGTGCAACTTATGACATGAAGTATGG 389

Qy 576 TGTTAGAGCTGACAGGGAGTGGAGGTCTCTCGGCGAGAAAACCGTTGGAAGGGTGGC 635
Db 388 TGTCAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGAGACCAATTGGAAGAGTTGC 329

Qy 636 TCGAGGGCCCTCCCAAGAAATTTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTT 695
Db 328 TCTCGAGCTTTGGGCAAGAGAATTTTGAAGCAATTTGCAAGCAATCTGAGAACTTTCGCTA 269

Qy 696 TGTTTCCAAAGTCATCAGTTGTGCTCCAGAGACGCGGTGATTTGATGCTGTGTAAC 755
Db 268 TGTCGCGCAGTTTCAACCATGTGTACTTCCAGAAAGATTTGTAGACAGAAATTTACAC 209

Qy 756 TTGTGAACAGATAGAGAGCAACATCGTTAGATGTCTCATGATCAGAGTACGAGAGAAGAT 815
Db 208 TCCG--ACAGATAGAAAATAACATTTGTGAGAGCCCTAAATCCGAAATAATCGAAAAGA 151

Qy 816 GATAGCCGATCGACAGATGAGGTTGAGGGGATTCGGTGGTGGAGTATCATCATG 875
Db 150 GATAACTCGCAATGAATCCGTCAGCAGATAGGAAATCTTGTGGTGGGTTTGGACCTG 91

Qy 876 CGTCCGTAGAAA 887
Db 90 CATTGTCGAAA 79

RESULT 8
US-10-267-763-7

; Sequence 7, Application US/10267763
; Patent No. 6800459
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1053)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-7

Query Match      16.9%; Score 274; DB 4; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6,1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGGTGAAGGCATCTGGAACACAGTTTGGAACTACTTTCAGGTTCGAACCTATGG 275
Db 306 CTTCCAGATACAAGCTACTGGAAGTTTCATATGGGACTCATTTTCGAGTTTCAACTTTGG 365

Qy 276 TGAATCTCATGGGGTGTGGTGTGGTATCATAGTGTGGTGTCCACTAGAAATCCACT 335
Db 366 AGAATCACAATGGAGAGGAGTTGGTTGTATCATTTGATGGTTGTCTCTCGTATTTCCACT 425

Qy 336 CACTGAGGCAGACCTACAAAGTTGAATCGATCGATGAGAGCGCCGCGACAGAGCAAAATAAC 395
Db 426 TACTGAAATCTGATTTGCAATTCGATCTCGATAGAGAGGCGCTGGTTCAGAGCAGGATCAC 485

Qy 396 CTCCACAAGGAAGGAGACTGATACATGCAAAATTTCTGTGAGGGACACATGAAGGGGTGAC 455
Db 486 AACTCCTAGAAAAGAGACTGATCTTGGCGGATATCTGTGGAGTCTCTGAGGAATGAC 545

Qy 456 TACTGGAACCCAAATCTTGTATTTGTGCCAAACACAGATCAAAATAGGCGAGTATCACCG 515
Db 546 GACAGGAACACCTATCCATGTGTTGTATCCAAACACAGATCAGAGAGGACTTGATTTACAG 605

Qy 516 TGAATAGCCAAATGTGTACCGACCTTCTCATGACAGCGCAACTTATGACTTCAAGTACGG 575
Db 606 TGAATATCGGTTTGGCTTATAGACCAATCGCATGCTGATGCAACTTATGACATGAAGTATGG 665

Qy 576 TGTTAGAGCTGTACAGGGAGTGGGAGGTCTCTCGGCGAGAAAACCGTTGGAAGGGTGGC 635
Db 666 TGTCAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGAGCCATTTGGAAGATTGC 725

Qy 636 TGCAAGGGCCCTCCCAAGAAAATTTTAAGCTCAAAATGTGGATTAAGATCTTGTGCTT 695
Db 726 TCCTGGAGCTTTGGGCAAGAGAATTTTGAAGCAATTTGCAAGCAACTGAGAAATCTTGCTA 785

Qy 696 TGTTTCCAAAGTCATCAGGTGTGCTCCAGAGAGCGCGTTGATTTATCGGTCTGTAAAC 755
Db 786 TGCTCGCAAGTTTCAACCAATGTGTACTTCCAGAGAAGATTTGTAGACACGAGAAATTTACAC 845

Qy 756 TTTGGAACAGATAGAGAGCAACATCGTTAGATGTCTGATCTCAGAGTACGAGAGAAGAT 815
```

Db 846 TCGG--ACAGATAGAAAATAACATTGTGAGAGCCCTAAATCCGAATAATCCGAAAAGA 903  
QY 816 GATAGACGCAATCGACAGAGTAGCAGTTTCGAGGGGATTCGGTCGGTGGAGTGCATCAGT 875  
Db 904 GATAACTCGCAATGAATCCGTCAGAGCATAGGAATACTTGTGTGGTGGTTTGACCTG 963  
QY 876 CGTCGCTAGAAA 887  
Db 964 CATTGTCGGA 975

RESULT 9  
US-09-248-796A-4099  
; Sequence 4099, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 4099  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (36)  
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd  
US-09-248-796A-4099

Query Match 12.5%; Score 203.8; DB 4; Length 1080;  
Best Local Similarity 52.1%; Pred. No. 4.1e-55;  
Matches 541; Conservative 0; Mismatches 473; Indels 24; Gaps 3;

QY 298 GGTGTGTATCAGTGGTGTGTCACCTAGAAATTCACCTCAGTCAGGAGCAGCTACAAGTT 357  
Db 7 GCTTTCTTCCCTGGTGGCAGAGTCGCGATGTCATTAACTAGAGCGATATTCACCA 66  
QY 358 GAATCGATCGAAGACGGCCCGACAGAGCAGAAATTAACCTCACAGGAAGAGACTGAT 417  
Db 67 CAATTGACTCGTATAGCTCCAGTCAAGTAAATTAATCTACCCCAAGAGATGAAAAAGAT 126  
QY 418 ACATGCAAAATTCGTCAAGGACACATGAAGGGTGACTACTGGAACGCCAATTCCTGTT 477  
Db 127 TTAGTTGAAATCAAAAGTGGAACTGAGAATGATTAACCTTGGGTTACCTATGGAATG 186  
QY 478 ATTGTCCTCAACACACAGATCAAAATAGGAGTGCAGTGCAGTGAATAGCCAAATGTACCCA 537  
Db 187 ATTGAGAAATAAAGATCATCGTCTGTTGATATAGTGAGCCGATTTGTTATCCAAGA 246  
QY 538 CTTCTCATGACAGCAACATTTATGACTTCAAGTACGGTGTGTAGAGCTGTACAGGAGGT 597  
Db 247 CCAAGTCATGCCGATGGACCTATATACAGAAATAGCGAAACCAATCTAGTTCTGGTGT 306  
QY 598 GSDAGTCTCTGGGAGAAAACCGTTGGAAGGGTGGCTGACAGGGCCCTCCCAAGAA 657  
Db 307 GGTAGATCATCTGCTAGAGAAACTATTTGGTAGAGTTGCCGACAGAGCCATTCAGAAAAG 366  
QY 658 ATTCTTAAGCTCAATGTGGATTAGAGATCTTGTGTTGTTTTCAAAGTGCATCAGGT- 716  
Db 367 ATTTTGGCCAAAGTGAATATGTTGAATTTGTCCTTTGTAAGTGTCTATTGTTGAATA 426  
QY 717 -----TGTGCTCCCAAGAACCGGGTTGATTATGGGTCTGTAACCTTTG 759  
Db 427 TCTATGAATAAATCTCCTCAAGATGCAAAATTTCCAAGAACTTTTAAACACTATCACTAGA 486

QY 760 GAACAGATAGAGAGCAACATCG---TTAGATGTCCTGATCCAGAGTACGACAGAGATG 816  
Db 487 GAACAAAGTCGATGGTGTAGTCCAAATAGATGTCAGATGCAAAATGTCGTGAGAAATG 546  
QY 817 ATAGACGCAATCGACAGAGTAGCAGTTTCGAGGGGATTCGGTCGGTGGAGTGCATCAGT 876  
Db 547 GTTAAAGTGAATTGAAAAATATCGTGATGCTAAAGATTCATTGGTGGTGTGTTCTCACTTGT 606  
QY 877 CTCGCTAGAAAAGCTTCCTCGGGGCTCGGTTCTCTCTGTTGTTGCAAGAGTCCAAATCCGAA 936  
Db 607 GTTATCGAAAACCTGTCCAATTGGATTAGGTGGAACCATGTTTCGATAAAATTTGGAAGCTAAA 666  
QY 937 CTGGCAAAAAGCTTATCTCTTCTATTCTCGAGCAACGGTTTCGAGATTGGCAGCGGATTC 996  
Db 667 TTGGCTCATGCAATGTTATCATACCCGCTACCAAGGGTTTGAATTTGGTCTCGATTT 726  
QY 997 GCCGGGACCGACTTCGACAGGAAGTAGCATATATGATGATT---TTATATGGATAAGGCT 1053  
Db 727 GAAGGTATCAAGATCCCTGGTTCAAAACATAACGATGCAATCTATATATGATGAAAAATTC 786  
QY 1054 GGAAGTGTGAGGACAGGAGCACTAATCGCTCGGGTGTGTGCGAGGGAGGATATCGATGTT 1113  
Db 787 GGAAGATTAAAGAACTGAAACAAACAATAGTGTGTTATCCAAAGGAGGAATTTTCAACCGT 846  
QY 1114 GAGATTGTGCACTTCAAAAGTTGCTTTTAAAGCCGACACCATCTATCGGGTGAACAGAAC 1173  
Db 847 GAAAACATTTATTTCTCAGTTGCCCTCAATCGGGTGTCTATCATGTCAGTAAGAACAGAA 906  
QY 1174 ACCGTGTCAAGGGAGCGTCAGAAAGCTTGAGCTTCTAGCAAGAGGCGCCATGACCCATGC 1233  
Db 907 ACTGCTACTACGATGGTAAAAGTGGTGTGTTGGCTGTAGAGGTAGACACGATCCAAGT 966  
QY 1234 GTGCCCCCTCGAGCTGTTCTGTGTGTAATCCTAGCGCGCTGTTGGTCTCATGACCG 1293  
Db 967 GTCAACCCCAAGAGCAGTTTCAATTTGTTGAAGCCATGACTGCAATAGTATTATGTGACGAA 1026  
QY 1294 CTGATGGCGCACGTGGCT 1311  
Db 1027 TATATGATTCAACAGCT 1044

RESULT 10  
US-09-313-294A-4734  
; Sequence 4734, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 4734  
; LENGTH: 300  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700348949H1  
; NAME/KEY: unsure  
; LOCATION: 44, 54  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-4734

Query Match 11.4%; Score 185; DB 4; Length 300;  
Best Local Similarity 84.3%; Pred. No. 2.1e-49;  
Matches 231; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 546 TGCAGACCAACTTATGACTTCAAGTACGGTTAGAGCTGTA-CAGGGAGGTGGAGGT 604

Db	1	TGCAGATGCAACCTATGACTTCAAGTATGGAGTTAGAGCTGTANCAGGAGACNGAAGGT	60
Qy	605	CCTCGGGCAGAAAAACCGTTGGAGGGTGGCTGCAGGGGCGCCCTCCCAAGAAAAATTCCTTA	664
Db	61	CATCAGCCAGAGAAACCAATTTGCGAGGGTGGCTGCAGGAGCTCTTCGAAAGAAAAATCTAA	120
Qy	665	AGCTCAAAATGTGGATTAGAGATCTTTGCTTTGTTTCCAAAGTGCATCAGGTGTGCTCC	724
Db	121	AGCTCAAAATCAGGAGTGCAGATCTTTGGCATTTGTTCTTAAAGTGCAACAGTCGTACTTC	180
Qy	725	CAGAAGACCGGTTGATTATGGCTCTGTAACTTTTGGAAACAGATAGAGAGCAACATCGTTA	784
Db	181	CAGAAGATGCGATTGATTATGAGACTGTAACTTGGAAAC-TATAGAGAGCAACATCGTTA	239
Qy	785	GATGTCCTGATCCAGAGTACGCAGAGAAAGATGAT	818
Db	240	GATGTCCTGATCCAGAAATATGCAGAGAAAGATGT	273
RESULT 11			
US-09-790-988-1/c			
; Sequence 1, Application US/09790988			
; Patent No. 6632935			
; GENERAL INFORMATION:			
; APPLICANT: SHIGENOBU, SHUJI			
; APPLICANT: WATANABE, HIDEMI			
; APPLICANT: HATTORI, WASHIRA			
; APPLICANT: SAKAKI, YOSHIYUKI			
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS			
; FILE REFERENCE: 081356/0159			
; CURRENT APPLICATION NUMBER: US/09/790,988			
; CURRENT FILING DATE: 2001-02-23			
; PRIOR APPLICATION NUMBER: JP2000-107160			
; PRIOR FILING DATE: 2000-04-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 640681			
; TYPE: DNA			
; ORGANISM: Buchnera sp.			
US-09-790-988-1			
Query Match 11.1%; Score 180.2; DB 4; Length 640681;			
Best Local Similarity 50.1%; Pred. No. 2.2e-45;			
Matches 532; Conservative 0; Mismatches 493; Indels 36; Gaps 2;			
Qy	232	TCTGAAACACGTTTGGAAACTACTTTTCAGGTTGCAACTATGTTGGAATCTCATGGGGT	291
Db	102985	TCTGAAATACAAATGGGAAATAATCTGTGTAAACCACTTTTGGCGAGTCACACGAGAA	102926
Qy	292	GGTCTTGGTTGTGTTATCAGTGGTTGTCCACTAGAAATCCACTCAGTGAGCAGACCTA	351
Db	102925	GCAATAGGTTGCATAAATTTGACGGAACACCTCCTCGTCTTGAATATCTTGTAAAGATTG	102866
Qy	352	CAAGTTTGAACCTCGATCGAAGACGGCCGACAGACAGAAATAACCTCCACAGGAAGAG	411
Db	102865	CAATATGATTTTGAATCTGTAGAGACAGGACTCTCCGTTTACACACTTTGGCGCGTGA	102806
Qy	412	ACTGATACATGCAGAAATTTCTGTGAGGACACATGAAGGGGTGACTACTGGAAACGCCAAT	471
Db	102805	CCTGATGAAGTTAATATATACTTTCCGGTATATTTAACGGGGTCAACAATGGTACTAGTAT	102746
Qy	472	CTTGTATTGTCCCAACACAGATCAATAGCAGTGATCACCGTGAATAGCCAAATGTG	531
Db	102745	GGTTTAATTAATTTAATATCATGATCATAGGCTCTCAAGACTATAGCGATATAAGAAATTTA	102686
Qy	532	TACGACCTCTCATGCAGACGCACTTATGACTTTCAAGTACGGTTGTAGAGCTGTACAG	591
Db	102685	TTTCGACCGGNCATGCCGATTATACCTTATGAAAAAATATGGAATTAGAATTTATCGC	102626
Qy	592	GGAGTGGAGGTCTCTCGGCGAGAAAAACCGTTGGAAAGGGTGGCTGCAGGGGCCCTCC	651



Db 8 CGGAAACAGTATCGGCAATTAATTAGAGTAACCACTTTTGGTGAGTCTCATCGGCACAG 67  
QY 293 GTGTTGGTGTGTTATCAGTGGTGTCCACTAGAAATTCACCTGAGCGACGACCTAC 352  
Db 68 CGTTAGGTTGATGTTGAGTGTCTCCCGGACTACCTTTAAACGAAGCGGATCTAC 127  
QY 353 AAGTTGAACCTCGATCGAAGACCGCCCGACAGACAGACAGAACTCCACCAAGGAAGGAGA 412  
Db 128 AAGTTGATTTAGATAGACGTAACCGGGAACCTTCACGTTATACACAAACGATAGAGC 187  
QY 413 CTGATCATGCAAAATTCCTGTGAGGACACATGAAGGGGTGACTACTGGAACGCCAATTC 472  
Db 188 CTGATCAAGTGCCTTTTATTCGGGTGTTTAAATGCTGTAAACAAAGGAACAGTATTG 247  
QY 473 TTGTTATTGTCCTCCAAACACAGATCAAAATGAGCAGTATCCCTGGAATAGCAATGTGT 532  
Db 248 GATTTAATAGAAATACCGATCAGCGCTCTCAAGATTAATAGCGAAATTAAGATGTAT 307  
QY 533 ACCGACCTTCTATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGG 592  
Db 308 TCCGCCCCAGGCGATGCAGATACACCTATGAACAGAAATATGTTACGTGATATCGTG 367  
QY 593 GAGTGGAGGCTCTCCGGCAGAAAACCGTTGGAAGGGTGTGAGGGGCCCTCCCA 652  
Db 368 GCGGGGACGCTCTTCTGTCGAGAAACCGCCATGCGTGTGCGAGCGGTGCTATTGCCA 427  
QY 653 AGAAATTTCTTAAGCTCAAAATGGAATTAGAGATCTTGCTTTGTTTCCAAAGTGCATC 712  
Db 428 AAAATACCTTAACAAAAATTTGTTATGAGTAAAGGTACTTATCTCAATT----- 482  
QY 713 AGTTGTGCTCCAGAAAGCGGGTTGATTATGGTCTGTAACTTTGGAACAGATAGAGA 772  
Db 483 -----AGGCGCGATTAGTTGTGAGTTAGTTGTTGATGTTGTTAGTAA 526  
QY 773 GCACATCGTTAGATGCTGATCCAGAGTACGACAGAGAGATGATAGCGCAATCGACA 832  
Db 527 CCACCCATCTTTTCCCGAGATCTCTCGCTAGATGCCCTTGATGAATATATCGGAG 586  
QY 833 GAGTACGAGTTCCAGGGGATTCGGTGGAGTGATCATGCGTCCGTAGCAAAAGCTTC 892  
Db 587 CGCTAAAAAAGAGGTAAATCTATAGTTGCCAAGTTACCGTGTGCGAAGGTGTAC 646  
QY 893 CTCGGGGTCTGTTCTCTGTGTCGACAACTCGAATCCGAATCGCAAAAGCTATGC 952  
Db 647 CTCGAGGATTAGTGAACCGTTTGTGATAGACTTGATGCCGATTTGGCTCATGCGTTAA 706  
QY 953 TTTCTATTCTCGAGCAGCGGTTGCGATTTGCGAGCGGATTCGCGGGGACCGACTTGA 1012  
Db 707 TGAGCATCAACGAGTAAAGCGGTTGAAATCGGAGATGGTTTGTATGTTGACCTTAA 766  
QY 1013 CAGGAAGTGAGCATAAATGATGAGTTTATATGATAAGGCTGGAAGTGTGAGGACACGA 1072  
Db 767 AAGGAGCTGAACCCCGATGAATTAACCAAGAGGATTTAGCAG----- 812  
QY 1073 CTAATCGCTCGGTTGTCAGGAGGATATCGAATTTGAGATTTGCACTTCAAAG 1132  
Db 813 -TAATCAACGCGGGGCGGTTATAGCGGATTTAGTGGTCAACTATTATTGACACA 871  
QY 1133 TTGCTTTTAAAGCGGACACCATCTATCGGGGTGAACAGACACCGTGTCAAGGGCGTC 1192  
Db 872 TTGCTTTTAAAGCGGACACTTCAAGTATTAATTTGCTGGTGAAGACACTTAATCGAAGGTG 931  
QY 1193 AGAAGTTGAGCTTCTAGCAAGAGCGGCCCATGACCCATGCGTCCGCTCGAGCTGTC 1252  
Db 932 AAGAGTGCATATGATCAACAAAGGCGACATGATCCGTTGTCGGTATCCGACTGTC 991  
QY 1253 CTGTTGTTGGAATCAATGCGCGGTTGTCCTCATGAGCACGCTGATGCGGCACTGGCTC 1312  
Db 992 CTAATGCAAGAGCTATGATGCTATTGTTTACTTGTATCATGCTGCTCAACGCGGTC 1051  
QY 1313 AGTGG 1318  
Db 1052 AATGTG 1057

RESULT 13  
US-08-916-421B-1  
; Sequence 1, Application US/08916421B  
; Patent No. 6503729  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
; Patent No. 6503729  
; FILE REFERENCE: PB275  
; CURRENT APPLICATION NUMBER: US/08/916,421B  
; CURRENT FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (103998)..(103998)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (148948)..(148948)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (163385)..(163385)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191989)..(191989)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191995)..(191995)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (231980)..(231980)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (234187)..(234187)





Db 357920 AGATTGGAAGATTGAAGGAGATTTTCAGCTACTATATAAATCCAGAGTTTGTGAAATG 357979  
QY 752 TAACTTTGGAACAGATAGAGCAACATCTTAGATGTCC-----TGATCCAG 799  
Db 357980 AAAAATCCTTAGAGATTAATAGAGATTATTGAAAGTATCCATTGAGATGCCATCAA 358039  
QY 800 AGTACCCAGAGAAGATGATAGACGCAATACACAGAGTACGAGTTGAGGGGATTCGGTCG 859  
Db 358040 TGAATGAGAAGAGATGGAGGAGTATGTTTAAAGCAATGGAATAAAGATAGTGTG 358099  
QY 860 GTGAGTATCATACGCTCGTAGAAAGTTCCTCGGGCTCGGTCTCCTGTTGTTCCG 919  
Db 358100 GAGGAGTTCTTGAATTTGTTGCAATTAATGTTCTGTTGGAGTTGGAATCCAATATCCA 358159  
QY 920 ACAAGCTCGAATCCGAATGCGCAAAAGTATGCTTTCTATTCTCGCAGCAACGGTTCCG 979  
Db 358160 ATAGTTTAAATGAGAAATTTGGCAGAGCTTTAATGAGTATAAATGCTGTTAAAGGAGTTG 358219  
QY 980 AGATTGGCAGCGATTTCGCGGACCGACTTTCACAGGAAGTGAACATAATGATGATTTT 1039  
Db 358220 AGATAGGGCTGTTTAAAGCGCTGAGATGTATGGAAGTGAAGTGAAGTGAAGTGT 358279  
QY 1040 ATATGATAGGCTGGAAGTGTTCAGACAGGACTAATCGCTCGGTGTGTGCGAGGAG 1099  
Db 358280 ATTTGATGACGACAAAAATATAAGATTCAAAACAAACCACTCGGTGTCATATTGGGAG 358339  
QY 1100 GGTATCGAATGTGAGATTGTGCACTTCAAAAGTGTCTTTAAGCGCACACCATCTATCG 1159  
Db 358340 GAATTAGCTGTGAACTCAATAGTTTAAAGAAATTCAGTAAGCCAAACACCTTCAATAG 358399  
QY 1160 GGGTGAACAGAACCCGT---GTCAAGGGAGCGTCAGAACGTTGAGCTTTCTAGCAAGAG 1216  
Db 358400 GTAAAAAGCAAAAACCAATAAATTTAAACCTTAGAAATGTTGAAATGAAATGAAG 358459  
QY 1217 GGGCCATGACCATCGCTGCGCCCTCGAGCTGTTCTGTGTGGAATCCATGCGCGCT 1276  
Db 358460 GAAGACACACCCAGTTATAGTTTCAAGGATTTCCAGTGGCTGAAGCAATGGTTGCTA 358519  
QY 1277 TGGTCTCATGACACCATGAT 1298  
Db 358520 TAACTTAGCTGATTTGATGAT 358541

RESULT 14  
US-09-692-570-1  
; Sequence 1, Application US/09692570  
; Patent No. 6797466  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii  
; Patent No. 6797466  
; TITLE OF INVENTION: jannaschii  
; FILE REFERENCE: PB275C1  
; CURRENT APPLICATION NUMBER: US/09/692,570  
; CURRENT FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; PRIOR APPLICATION NUMBER: US 08/916,421  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g

; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (103998)..(103998)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (148948)..(148948)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (163385)..(163385)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (191989)..(191989)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (191995)..(191995)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (231980)..(231980)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234187)..(234187)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234220)..(234220)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234814)..(234814)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (309398)..(309398)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600592)..(600592)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Query Match 10.9%; Score 178; DB 4; Length 1664976;
Best Local Similarity 49.9%; Pred. No. 2.3e-44;
Matches 550; Conservative 0; Mismatches 525; Indels 27; Gaps 3;

Qy 224 TGAAGGCATCTGGAACACGCTTTGGAACTACTTTCAGGTTTGCACCTTAGAATTCCTACTCTGAGG 283
Db 357440 TGGTGACATTAATGAACACCTTATGGGGATATGTTTAGAGTTTACAGTTTGGAGAAAGTC 357499

Qy 284 ATGGGGTGGTGTGGTGTGTTTATCAGTGTGTTCCACCTTAGAATTCCTACTCTGAGG 343
Db 357500 ATGGAAGGCTGTTGGAGCAGTTGTTGATGGATGTCAGCTAATCTGCTTTTATCTGAAG 357559

Qy 344 CAGACCTACAAGTTGAATCTGATCGAAGAGCGGCCGACAGAGAGAGAGAGATTAACCTCCACAA 403
Db 357560 AGGATATCCAAAAGAGCTTGCACAGGAGAAAGACCAGGGCAGAGCATCTTCTCAACACCAA 357619

Qy 404 GGAAGGAGACTGTATACATGCAAAATTTCTGTCCAGGACACATGAAGGGTGACTACTGAA 463
Db 357620 GAAAAGAGAGATAAAGTTGAATCTTATCAGAAATTTTGGGGGAAAATCTACTGAG 357679

Qy 464 CGCCAAATCTTGTATTGTCTCCAAACACAGATCAATAGGCGATGATCACCGTGAATAG 523
Db 357680 CTCCTATTGCTCAATAGTCTATAACAAAACATGAGACCTTAAAGATTACTCAAAAATTA 357739

Qy 524 CCAATGTGTACCGACCTTCTCATGCAGACGCACTTATGACTTCAAGTACGGTGTAGAG 583
Db 357740 AAGTACACCAAGACCTGGACATGCAGATTTAACTATAGATTGAAGTATATAAAACTATG 357799

Qy 584 CTGTACAGGGAGGTGGGAGGTCCTCGGGCAGAAAACCGTTGGAAGGTCGCTGCAGGG 643
Db 357800 ATTATAGGGAGGAGGAGGCAAGTGTAGAGTAAAGTAAAGTATGTTATTGGAGGAG 357859

Qy 644 CCCTCCCCAAGAAATTTCTTAAGCTCAAATGTGGATTAGAGATCTTGTGCTTTGTTCCA 703
Db 357860 CTATTGCTAAAAGCTTCTATCTTACACATACACATAAAAATTTATTGGTTATACCATAA 357919

Qy 704 AAGTGCATCAGGTTGTG-----CTCCCAAGAGACCGGTTGATTATGGGTCG 751
Db 357920 AGATTGGAAGATTGAAGGGAGATTTTCAGCTACTATAAAAATCCAGAGGTTTTTCAAAATG 357979

Qy 752 TAACCTTTGGAACAGATAGAGCAACATCGTTAGATGTC-----TGATCCAG 799
Db 357980 AAAAATCCTTAGAGAGATTATAGAGATTATTGAAGTAAATCCATTGAGATGTCATCAA 358039

Qy 800 AGTACGACAGAGAGATGATAGACCGCAATCGACAGATTCGAGTTTCGAGGGGATTCGGTCG 859
```

Db 358040 TGAATGAGAAAGAGATGAGGAGTATGTTTTAAAGGCAATGGAATAAAGATAGTTG 358099  
Qy 860 GTGAGATGATCACATGCTGCTAGAAACGTTCTTCGGGCTCGTTCCTGTTGTCG 919  
Db 358100 GAGAGTTGTTGAATTTGTCATTAATGTTCTGTTGGAGTTTGGAAATCCCAATATCA 358159  
Qy 920 ACAAGCTCGAATCCGAACCTGCAAAAGCTATGCTTTCTATTCTCTGCGAGCAACGGGTTG 979  
Db 358160 ATAAGTTAAATGGAGAATTGGCAGAGCTTTAATGATATAAATGCTGTTAAAGGAGTTG 358219  
Qy 980 AGATTGGCAGCGGATTCGCGGGACCGACTTGCAGAGAGTGAGCAATAATGATGATTTT 1039  
Db 358220 AGATAGGGCTGTTGTTTTAAAGCGCTGAGATGATGAAAGTGATGAGTGAACGATGAGATG 358279  
Qy 1040 ATATGGATAAGGCTGGAGTGTGAGGACCGGACTATCCCTCGGTGGTGTGCGAGGAG 1099  
Db 358280 ATTTGATGACGCAAAATAATAGATTCAAAACAAACAACTCGGTGGGATATTGGGAG 358339  
Qy 1100 GGAATATCGAATGTTGAGATTGTGCACCTTCAAAGTTGCTTTTAAAGCCGACACCACTATCG 1159  
Db 358340 GAATTAGCTGTGAACCTCCAAATAGTTTAAAGATTGCGATTAAAGCCAAACCTTCAATAG 358399  
Qy 1160 GGGTGAACAGAACACCGT---GTCAAGGAGCGTCAGAAAGTTGAGCTTCTAGCAAGAG 1216  
Db 358400 GTAAAGAGCAAAACCAATAAATTTAAACACCTTAGAAAATGTTGAAATTTGAAATTTGAAG 358459  
Qy 1217 GCGCCATGACCCATGCTGCTGCGCCCTGAGCTGCTCTGTTGGTGAATCCATGGCCGCT 1276  
Db 358460 GAAGACAGGACCGATTATGTTTCAAGGATTGTTCCAGTGGCTGAAGCAATGGTTGCTA 358519  
Qy 1277 TGTCTCTCATGACACGCTGAT 1298  
Db 358520 TAACCTTAGCTGATTGATGAT 358541

## RESULT 15

US-09-064-693A-20  
; Sequence 20, Application US/09064693A  
; Patent No. 6210937  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Thomas E.  
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY  
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION  
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: W. Gary Goodson  
; ADDRESSEE: INEEL--Lockheed Martin Idaho  
; ADDRESSEE: Technologies Co.  
; STREET: P.O. Box 1625  
; CITY: Idaho Falls  
; STATE: Idaho  
; COUNTRY: USA  
; ZIP: 83415-3810  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
; COMPUTER: Toshiba Satellite Pro T2150CDS  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: Word Perfect 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/064,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: W. Gary Goodson  
; REGISTRATION NUMBER: 22,387  
; REFERENCE/DOCKET NUMBER: Lit-PI-296  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (208)526-9469  
; TELEFAX: (208)526-8339

## ; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1239 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-09-064-693A-20

Query Match 10.5%; Score 170; DB 3; Length 1239;  
Best Local Similarity 49.9%; Pred. No. 4.5e-44;

Matches 540; Conservative 0; Mismatches 505; Indels 37; Gaps 3;

Qy 233 CTGGAACACGTTTGGAACTTCTTTCAGTTTGCACCTATGTTGTAATCTCATGCGGGTG 292  
Db 128 CTGGAACACCAATTGGACAACTCTTTTCGCTAAACCACTTCGGGAATCGCACGCGCTG 187  
Qy 293 GTGTTGGTGTGTTATCATGTTGTTCTCCACCTAGAAATCCACTCACTCAGGAGACCTAC 352  
Db 188 CGCTCGGCTGCATCGTTCGATGTTTCCGAGGCAATTCGCTGACGGAAGCGGACCTGC 247  
Qy 353 AAGTTGAACTCGATCGAAGACGGCCGACAGAGCAGAAATAAACCCTCCACAAAGGAGAGA 412  
Db 248 AACATGACTCGACCGCTGCTGCTCGCTGGGACATCGGCTATACCAACCAGCGCCGAGC 307  
Qy 413 CTGATACATGCAAAATTTCTGTGAGGGAACATGAAGGGGTGACTACTGGAACGCCAATTC 472  
Db 308 CGGATCAGGTCAAAATTTCTCTCGGTGTTTGAAGGGGTTACTACCGGCACCGACATTCG 367  
Qy 473 TTGTTATTGTTCCCAACACACATCAATAGCAGTGATCAGCGTGATCAGCGTAATACCCATGTGT 532  
Db 368 GTTGTGATGCAAAACACTGACAGGCTCTCAGGATTTACAGTGCGATTAAAGACGTTT 427  
Qy 533 ACCGACCTTCTCATGACAGACGAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGG 592  
Db 428 TCCGTCCAGGCAATGCCGATTACACTACGAACAAATACGGTCTGCGGATTTATCGCG 487  
Qy 593 GAGGTGGGAGTCTCTCGGCGAAGAAACCGTTGGAAGGGTGGCTGCAAGGGGCCCTCCCCA 652  
Db 488 GCGGTGGAAGCTTCTTCCGCGCGGAAACCGGCATGCGCGTGGCGGAGGAGCTATTGCCA 547  
Qy 653 AGAAATTTCTTAAGCTCAAAATGTGGAATTAGAGATCTTGTCTGTTGTTTCCAAAGTGATC 712  
Db 548 AAAATATCTCGCGGAAATTTGTAATGAAATTCCTGGCTGCTGACCCAGATGGGCG 607  
Qy 713 AGTTGTGCTCCCAAGAGACGCGTGTGATTATGGGTCTGTAACCTTTGGAACACAGATAGA 772  
Db 608 ACATTCCGCTGGATATCAAAAGC-----TGTCTCAGGTTCAGC 646  
Qy 773 GCAACATCGTTAGATGCTCTGATCCAGAGTACGAGAGAAAGATGATAGACGCAATCGACA 832  
Db 647 AAATCCGTTTTCGCGGACCCGACAAATTCGACGCGTTAGACGAGTTGATGCGTG 706  
Qy 833 GAGTACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTGATCATCGCTCGCTAGAAACGTTTC 892  
Db 707 CGCTGAAAAGAGAGGGGCACTCCATCGGCGCTAAAGTCAACGTTGTTCCAGTGGCGTTC 766  
Qy 893 CTGCGGGCTCGGTTCTCTGTGTTTCGAAAGCTCGAATCCGAATCGGCAAAAGCTATGC 952  
Db 767 CTGCGGACTTGGGAGCGCGTCTTTGACCGCTCGGATGCTGACATCGCCCATGCGCTGA 826  
Qy 953 TTTCTATTCGCGAGCAACGGGTTTCGAGATGGCAGCGGATTCGCGGGACCGACTTGA 1012  
Db 827 TGAGCATCAACGCGGTGAAAGCGCTGGAATTTGGCGACGCTTTGACGTTGGTGGCGCTGC 886  
Qy 1013 CAGGAAGTGAGCATAATGATGATTTTATATGATAAGCTGGAAGTGTTCAGGACACGGA 1072  
Db 887 GCGGACGACGAAACCGGATGAAATCACCAGA-----CGTTTCACA 931  
Qy 1073 CTAATCGCTCGGTTGTGTGCGAGGAGGATATCGAATGTTGAGATTTGCACTTCAAG 1132  
Db 932 GCAACCATGCGGCGGCAATTTCTCGCGGTATCAGCAGCGGCGAGCAATCAATTGCCATA 991  
Qy 1133 TTGCTTTTAAAGCGCACACCATCTATCGGGGTGAAACAGAACACCGTGTCAAGGAGCGTC 1192

Db	992	TGGCGCTGAAACCGACCTCCAGCATTACCGTGCCGGGTCGTACCATTAACCGCTTTGGCG	1051
Qy	1193	AGAACGTTGAGCTTCTAGCAAGAGGGCGCCCATGACCCATGCGTCGCCCTCGAGCTGTC	1252
Db	1052	AAGAACTTGAGATGATCACCAAGGCCGTCACGATCCCTGTGTGGGATCCGCGCAGTGC	1111
Qy	1253	CTGTGGTGAATC-CATGGCCCGTTGGTCTCTCATGGACCAAGTGGCGCACGTGGCT	1311
Db	1112	CGATCGCAAGACGGAATCGTGGCGATCGTTTTAATGGATCACCTGTTACGGCAACGGGCG	1171
Qy	1312	CA	1313
Db	1172	CA	1173

Search completed: August 27, 2005, 04:27:18  
Job time : 318 secs

GenCore version 5.1.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 00:49:04 ; Search time 1018 Seconds  
 (without alignments)  
 10450.947 Million cell updates/sec

Title: US-10-677-179-7

Perfect score: 1626

Sequence: 1 gcaagagctcagcttcgtctt.....caaaaaaaaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:  
 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:  
 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:  
 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:  
 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	1626	18	US-10-677-179-7
2	1385	85.2	1496	18	Sequence 7, Appli
3	1382.2	85.0	1541	20	Sequence 30604, A
4	1321.8	81.3	1845	20	Sequence 138335
5	872	53.6	1792	20	Sequence 138337
6	870.4	53.5	1635	18	US-10-425-115-171490
7	867.2	53.3	1603	18	US-10-677-179-1
					Sequence 1, Appli
					Sequence 32244, A

#### ALIGNMENTS

##### RESULT 1

US-10-677-179-7  
 ; Sequence 7, Application US/10677179  
 ; Publication No. US20040082050A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cahoon, Rebecca E.  
 ; APPLICANT: Falco, Saverio Carl  
 ; APPLICANT: Pember, Stephen O.  
 ; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
 ; FILE REFERENCE: BB-1159-A  
 ; CURRENT APPLICATION NUMBER: US/10/677,179  
 ; CURRENT FILING DATE: 2003-10-02  
 ; PRIOR APPLICATION NUMBER: US/09/743,207  
 ; PRIOR FILING DATE: 2001-01-04  
 ; PRIOR APPLICATION NUMBER: PCT/US99/16353  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/093,611  
 ; PRIOR FILING DATE: 1998-07-21  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 7  
 ; LENGTH: 1626  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 US-10-677-179-7

Query Match 100.0%; Score 1626; DB 18; Length 1626;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 99207, A  
 Sequence 11, Appl  
 Sequence 8882, Ap  
 Sequence 1811, Ap  
 Sequence 27, Appl  
 Sequence 5, Appli  
 Sequence 171487,  
 Sequence 13775, A  
 Sequence 9032, Ap  
 Sequence 13579, A  
 Sequence 43897, A  
 Sequence 42555, A  
 Sequence 26364, A  
 Sequence 1812, Ap  
 Sequence 9009, Ap  
 Sequence 8754, Ap  
 Sequence 46069, A  
 Sequence 44740, A  
 Sequence 1, Appli  
 Sequence 7, Appli  
 Sequence 9218, A  
 Sequence 45609, A  
 Sequence 34106, A  
 Sequence 9882, Ap  
 Sequence 12647, A  
 Sequence 6185, Ap  
 Sequence 1892, Ap  
 Sequence 1892, Ap  
 Sequence 31517, A  
 Sequence 41598, A  
 Sequence 41210, A  
 Sequence 24843, A  
 Sequence 42063, A  
 Sequence 6135, Ap  
 Sequence 47170, A  
 Sequence 20439, A  
 Sequence 5, Appli

Qy 1 GCACGAGCTCAGCTTCGTCTCTCTCGCCGGCGGCGAGCAACTATCATCACTTCATTAG 60  
Db 1 GCACGAGCTCAGCTTCGTCTCTCTCGCCGGCGGCGAGCAACTATCATCACTTCATTAG 60  
Qy 61 CTATCCAACTTATTCGATGACGACCGTGCCTCAAGCCACAGCGAGTGGCCACTCAAG 120  
Db 61 CTATCCAACTTATTCGATGACGACCGTGCCTCAAGCCACAGCGAGTGGCCACTCAAG 120  
Qy 121 GCACGGCTCGCACCCCGCGCGATCGCGCCCTTCTGCGAGTTTGCCTCAGCTTCCTCC 180  
Db 121 GCACGGCTCGCACCCCGCGCGATCGCGCCCTTCTGCGAGTTTGCCTCAGCTTCCTCC 180  
Qy 181 CTCGCTTCGCGCTGCACCGCTGCCTGCGCACTGCTGCGCTAGAGTGAAGCACTCGGAAC 240  
Db 181 CTCGCTTCGCGCTGCACCGCTGCCTGCGCACTGCTGCGCTAGAGTGAAGCACTCGGAAC 240  
Qy 241 ACGTTTGGAACTACTTTCAAGTTGCAACCTATGCTGAATCTCATGGGGGTGGTTGGT 300  
Db 241 ACGTTTGGAACTACTTTCAAGTTGCAACCTATGCTGAATCTCATGGGGGTGGTTGGT 300  
Qy 301 TGTGTTATCAGTGGTTGTCACCTAGAAATTCACCTCACTGAGGCGAGACTACAGTTGAA 360  
Db 301 TGTGTTATCAGTGGTTGTCACCTAGAAATTCACCTCACTGAGGCGAGACTACAGTTGAA 360  
Qy 361 CTCGATCGAAGACGGCCCGGACAGACAGATAAATCTCCACAGGAAGAGACTGATACA 420  
Db 361 CTCGATCGAAGACGGCCCGGACAGACAGATAAATCTCCACAGGAAGAGACTGATACA 420  
Qy 421 TGCATAATTCGTTCAGGACACATGAAGGGGTGACTACTGGAACGCCAAATTCCTGTTAT 480  
Db 421 TGCATAATTCGTTCAGGACACATGAAGGGGTGACTACTGGAACGCCAAATTCCTGTTAT 480  
Qy 481 GTCCCAACACAGATCAATAGGACGATACCGTGAAATAGCCAAATGTGTACCGACCT 540  
Db 481 GTCCCAACACAGATCAATAGGACGATACCGTGAAATAGCCAAATGTGTACCGACCT 540  
Qy 541 TCTCATGACAGACCAACTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGG 600  
Db 541 TCTCATGACAGACCAACTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGG 600  
Qy 601 AGGTCCTCGGCGAGAAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAT 660  
Db 601 AGGTCCTCGGCGAGAAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAT 660  
Qy 661 CTTAAGCTCAATATGGATTAGAGATCTTGTTGTTTCCAAAGTCATCAGTTGTG 720  
Db 661 CTTAAGCTCAATATGGATTAGAGATCTTGTTGTTTCCAAAGTCATCAGTTGTG 720  
Qy 721 CTCCCAAGACGCGGTTGATTATGGGTCTGTAACTTTGGAAACAGATAGAGCAACATC 780  
Db 721 CTCCCAAGACGCGGTTGATTATGGGTCTGTAACTTTGGAAACAGATAGAGCAACATC 780  
Qy 781 GTTAGATGTCCTGATCCAGAGTACGACAGAGATGATAGCGCAATCGACAGAGTACGA 840  
Db 781 GTTAGATGTCCTGATCCAGAGTACGACAGAGATGATAGCGCAATCGACAGAGTACGA 840  
Qy 841 GTTCAGGGGATTCGGTCGGTGGAGTGATCATCGCTCGTAGAAAGTTCTTCGGGG 900  
Db 841 GTTCAGGGGATTCGGTCGGTGGAGTGATCATCGCTCGTAGAAAGTTCTTCGGGG 900  
Qy 901 CTCGGTTCTCTGTGTCGACAAAGCTCGAAATCCGAACTGGCAAAAGCTATGCTTCTATT 960  
Db 901 CTCGGTTCTCTGTGTCGACAAAGCTCGAAATCCGAACTGGCAAAAGCTATGCTTCTATT 960  
Qy 961 CTTGCGAGCAACGGGTTTCAGATTCGCGAGCGGATTCGCGGGACCGAATTCACAGGAAGT 1020  
Db 961 CTTGCGAGCAACGGGTTTCAGATTCGCGAGCGGATTCGCGGGACCGAATTCACAGGAAGT 1020  
Qy 1021 GAGCATATGATGATGTTTATATGATGAAGCTGGAAGTGCAGACGAGCTAATCGC 1080  
Db 1021 GAGCATATGATGATGTTTATATGATGAAGCTGGAAGTGCAGACGAGCTAATCGC 1080  
Qy 1081 TCGGGTGGTGTGACGGAGGGATATCGAAATGTTGAGATTGTGCACTTCAAAGTTGCTTTT 1140

Db 1081 TCGGGTGGTGTGACGGAGGGATATCGAAATGTTGAGATTGTGCACTTCAAAGTTGCTTT 1140  
Qy 1141 AAGCCGACACATCTATCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT 1200  
Db 1141 AAGCCGACACATCTATCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT 1200  
Qy 1201 GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTGGTCCCTCGAGCTGTTCTGTGGTG 1260  
Db 1201 GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTGGTCCCTCGAGCTGTTCTGTGGTG 1260  
Qy 1261 GAATTCATGCGCGCTTGGTCTCATGACACAGCTGATGGCGCACGCTGAGTGGCGAG 1320  
Db 1261 GAATTCATGCGCGCTTGGTCTCATGACACAGCTGATGGCGCACGCTGAGTGGCGAG 1320  
Qy 1321 ATGTTCCGGCTCAATATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCAGC 1380  
Db 1321 ATGTTCCGGCTCAATATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCAGC 1380  
Qy 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT 1440  
Db 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT 1440  
Qy 1441 GGCTTGGTTTGTCTCTTTCACCGTAGTTTGTGTTTTTTTCCCGCAAGTGTGATG 1500  
Db 1441 GGCTTGGTTTGTCTCTTTCACCGTAGTTTGTGTTTTTTTCCCGCAAGTGTGATG 1500  
Qy 1501 CGATGAAGTGAATAAGGACACTTGGTTTCTGTGCACTTTGACAGTTTCAATAATGTA 1560  
Db 1501 CGATGAAGTGAATAAGGACACTTGGTTTCTGTGCACTTTGACAGTTTCAATAATGTA 1560  
Qy 1561 TCTACTTCGAAGATGATGATGATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620  
Db 1561 TCTACTTCGAAGATGATGATGATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620  
Qy 1621 AAAAAA 1626  
Db 1621 AAAAAA 1626

RESULT 2

US-10-425-114-30604  
; Sequence 30604, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 30604  
; LENGTH: 1496  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73045G02\_FLI  
US-10-425-114-30604

Query Match 85.2%; Score 1385; DB 18; Length 1496;  
Best Local Similarity 94.5%; Pred. No. 0;  
Matches 1490; Conservative 0; Mismatches 5; Indels 82; Gaps 2;  
Qy 12 GCTTCGCTCTCTCGCGCGCGCAGGCAACTATCATCTCATCTTATAGTTCATCCATC 71  
Db 1 GCTTCGCTCTCTCTCGCGCGCGCAGGCAACTATCATCTCATCTTATAGTTCATCCATC 60



QY 72 TATTTCGATGACGACCGTCCCAAGCCACAGCAGAGTGGCGCACTCAACGGGCGCGCTCGC 131  
 Db 61 TATTTCGATGACGACCGTCCCAAGCCACAGCAGAGTGGCGCACTCAACGGGCGCGCTCGC 120  
 QY 132 ACCCGCGCGATCGCGGCTTGTGAGTTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 191  
 Db 121 ACCCGCGCGATCGCGGCTTGTGAGTTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 QY 192 CGTGACCGCTGCGCGCACTGCTCGCTAGAGTGAAGGCACTTGGAAACACAGTTTGGAAA 251  
 Db 181 CGTGACCGCTGCGCGCACTGCTCGCTAGAGTGAAGGCACTTGGAAACACAGTTTGGAAA 240  
 QY 252 CTACTTTTCAGGTGCAACCTATGTTGAATCTCATGGGGTGGTGGTGGTGGTGGTGGTGGTGG 311  
 Db 241 CTACTTTTCAGGTGCAACCTATGTTGAATCTCATGGGGTGGTGGTGGTGGTGGTGGTGGTGG 300  
 QY 312 TGGTTGTCCACCTAGCAATTCCTACTGAGGCGAGACCTACAGTTGAACTCGATCGAAG 371  
 Db 301 TGGTTGTCCACCTAGCAATTCCTACTGAGGCGAGACCTACAGTTGAACTCGATCGAAG 360  
 QY 372 ACAGCGCGGACAGAGCAGAAATTAACCTCCCAAGGAGGAGACTGATACATGCAAAATCT 431  
 Db 361 ACAGCGCGGACAGAGCAGAAATTAACCTCCCAAGGAGGAGACTGATACATGCAAAATCT 420  
 QY 432 GTACGGGACACATGAAGGGGTGACTACTGGAACGCCAATTCCTGTTATTTGTCCTCCAAACAC 491  
 Db 421 GTACGGGACACATGAAGGGGTGACTACTGGAACGCCAATTCCTGTTATTTGTCCTCCAAACAC 480  
 QY 492 AGATCAATAGGCGAGTATCAACGTGAATAGCCAAATGATGTAACCGACTTCTCATGCGAGA 551  
 Db 481 AGATCAATAGGCGAGTATCAACGTGAATAGCCAAATGATGTAACCGACTTCTCATGCGAGA 540  
 QY 552 CGCAACTATGACTTCAAGTACGTTGAGTGTAGAGCTGTACAGGAGGTGGAGGTCTCGGG 611  
 Db 541 CGCAACTATGACTTCAAGTACGTTGAGTGTAGAGCTGTACAGGAGGTGGAGGTCTCGGG 600  
 QY 612 CAGAAAAACCGTTGGAAGGGTGGCTGCGAGGGGCTTCCCAAGAAAAATTTCTTAAGCTCAA 671  
 Db 601 CAGAAAAACCGTTGGAAGGGTGGCTGCGAGGGGCTTCCCAAGAAAAATTTCTTAAGCTCAA 660  
 QY 672 ATGTGATTAAGATCTTGTCTGTTGTTTTCAAAGTGCAATAGGTTGCTCCCGAAGA 731  
 Db 661 ATGTGATTAAG-----AGATAGAGAGCAACATCGTTAGATGCC 671  
 QY 732 CGCGGTTGATATGGTCTGTAACTTTGGAAACAGATAGAGCAACATCGTTAGATGCC 791  
 Db 672 -----AGATAGAGAGCAACATCGTTAGATGCC 699  
 QY 792 TGATCCAGAGTACGAGAGAGATGATAGACCAATCGACAGTACGAGTTCCGAGGGGA 851  
 Db 700 TGATCCAGAGTACGAGAGAGATGATAGACCAATCGACAGTACGAGTTCCGAGGGGA 759  
 QY 852 TTGCGTGGTGGAGTGATCAGATCGGTGCTAGAAAACGTTCTCGCGGCTCGGTTCTCC 911  
 Db 760 TTGCGTGGTGGAGTGATCAGATCGGTGCTAGAAAACGTTCTCGCGGCTCGGTTCTCC 819  
 QY 912 TGTGTTCCAGAGCTCGAATCGAATCGCAAGCTATGCTTTCTATTCCTCGGAGCA 971  
 Db 820 TGTGTTCCAGAGCTCGAATCGAATCGCAAGCTATGCTTTCTATTCCTCGGAGCA 879  
 QY 972 CGGTTTCAGAGTTGCGAGCGAATTCGCGGGACCGACTTTCAGAGAGTGAACATAATGA 1031  
 Db 880 CGGTTTCAGAGTTGCGAGCGAATTCGCGGGACCGACTTTCAGAGAGTGAACATAATGA 939  
 QY 1032 TGAGTTTATATGGAATAAGCTTGAAGTGTGAGGACAGCACTAATCGCTCGGGTGGT 1091  
 Db 940 TGAGTTTATATGGAATAAGCTTGAAGTGTGAGGACAGCACTAATCGCTCGGGTGGT 999  
 QY 1092 GCAGGGAGGATATCGAATGTTGAGATGTTGCACTTCAAAGTGTCTTTTAAGCCGACACC 1151  
 Db 1000 GCAGGGAGGATATCGAATGTTGAGATGTTGCACTTCAAAGTGTCTTTTAAGCCGACACC 1059  
 QY 1152 ATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGGAGGCTCAGAACGTTGAGCTTCTAGC 1211

Db 1060 ATCTATCGGGGTGAAACAGAACACCGTGTCAAGGAGCGTCAGAACGTTGAGCTTCTAGC 1119  
 QY 1212 AAGAGGCGGCATGACCCATGCGGTCGCCCTCGAGCTGTTCTGTGGTGGAAATCCATGCG 1271  
 Db 1120 AAGAGGCGGCATGACCCATGCGGTCGCCCTCGAGCTGTTCTGTGGTGGAAATCCATGCG 1179  
 QY 1272 CGCGTTGCTCTCATGAGCACGATGCGCGCACGTGGCTCAGTCCGAGATGTTCCGCGCT 1331  
 Db 1180 CGCGTTGCTCTCATGAGCACGATGCGCGCACGTGGCTCAGTCCGAGATGTTCCGCGCT 1239  
 QY 1332 CAATACTCCACTTCAAGAACCAAGTTGGCTCTTTTAGCAGAGGCGAGACACCTGATGA 1391  
 Db 1240 CAATACTCCACTTCAAGAACCAAGTTGGCTCTTTTAGCAGAGGCGAGACACCTGATGA 1299  
 QY 1392 GCTCGGCGCAATTTTATCATATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTG 1451  
 Db 1300 GCTCGGCGCAATTTTATCATATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTG 1359  
 QY 1452 CTGTGCTCTGCGACCGTAGTTTGTG-----TTTTTTTTTCCCGCAAGTGTGATCGATCAAGTG 1510  
 Db 1360 CTGTGCTCTGCGACCGTAGTTTGTGTTTTTTTTTCCCGCAAGTGTGATCGATCAAGTG 1419  
 QY 1511 AATAAGGCATTTGGTTCTCTGTCATTTGTACAGTTTTCATATATATGTAATCTACTTGA 1570  
 Db 1420 AATAAGGCATTTGGTTCTCTGTCATTTGTACAGTTTTCATATATATGTAATCTACTTGA 1479  
 QY 1571 AGATGATGATTTTAT 1587  
 Db 1480 AGATGATGATTTTAT 1496

RESULT 3

US-10-425-115-138335  
 ; Sequence 138335, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 138335  
 ; LENGTH: 1541  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURES:  
 ; OTHER INFORMATION: Clone ID: MFT4577\_57639C.1  
 US-10-425-115-138335

Query Match 85.0%; Score 1382.2; DB 20; Length 1541;  
 Best Local Similarity 93.2%; Pred. No. 0;  
 Matches 1501; Conservative 0; Mismatches 28; Indels 82; Gaps 2;  
 QY 12 GCTTCGTCTCTCTCGCGCGCGCGGCAACTATCATCACTTCAATTAGCTCATCAATC 71  
 Db 1 GCTTCGTCTCTCTCGCGCGCGCGGCAACTATCATCACTTCAATTAGCTCATCAATC 60  
 QY 72 TATTCCGATGACGACCGTGGCCCAAGCCACAGAGTGGCGCACTCACGGGCAACGGCTCGC 131  
 Db 61 TATTCCGATGACGACCGTGGCCCAAGCCACAGAGTGGCGCACTCACGGGCAACGGCTCGC 120  
 QY 132 ACCCGCGCGATCGCGGCTTGTGGAGTTTGGCCAGCTTCTCTCCCTCCGCTCGGTTGCG 191  
 Db 121 ACCCGCGCGATCGCGGCTTGTGGAGTTTGGCCAGCTTCTCTCCCTCCGCTCGGTTGCG 180  
 QY 192 CGTGACCGCTGCGCGCACTGCTCGCTAGAGTGAAGGCACTCGGAACACAGTTTGGAAA 251

```
Db 181 CGTGACCGCTGCCCACTGCTCGCTAGAGGTGAAGCATCTGGAACACGTTTGGAAA 240
Qy 252 CTACTTTTCAGGTTGCAACTGATGAGTGAATCTCATGCGGGGTGGTGTGTTGTGTTATCAG 311
Db 241 CTACTTTTCAGGTTGCAACTGATGAGTGAATCTCATGCGGGGTGGTGTGTTGTGTTATCAG 300
Qy 312 TGGTTGTCCACTAGAAATTCACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 371
Db 301 TGGTTGTCCACTAGAAATTCACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 360
Qy 372 ACGGCCCGGACAGACAGACAGATAACCTCCACAGGAAGGAGACTGATACATGCAAAATTCCT 431
Db 361 ACGGCCCGGACAGACAGACAGATAACCTCCACAGGAAGGAGACTGATACATGCAAAATTCCT 420
Qy 432 GTCAGGGAACATGAAAGGGGTGACTACTGGAACCGCAATTCCTGTTATGTTCCCAACAC 491
Db 421 GTCAGGGAACATGAAAGGGGTGACTACTGGAACCGCAATTCCTGTTATGTTCCCAACAC 480
Qy 492 AGATCAAAATAGCCAGTGATCCCGTGAATAGCCAAATGTAACGACCTTCTCATGCAGA 551
Db 481 AGATCAAAATAGCCAGTGATCCCGTGAATAGCCAAATGTAACGACCTTCTCATGCAGA 540
Qy 552 CGCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGAGAGGTGGAGGTCTCTCGGG 611
Db 541 CGCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGAGAGGTGGAGGTCTCTCGGG 600
Qy 612 CAGAAAAACCGTTGGAAGGGGTGCTGAGAGGGCCCTCCCAAGAAAAATCTTAAGCTCAA 671
Db 601 CAGAAAAACCGTTGGAAGGGGTGCTGAGAGGGCCCTCCCAAGAAAAATCTTAAGCTCAA 660
Qy 672 ATGTGGATTTAGAGATCTTGTGCTGTTGTTTCCAAAGTGCAATCAGGTGTGCTCCCAAGA 731
Db 661 ATGTGGATTTAG----- 671
Qy 732 CGCGGTTGATTTATGAGGTCTGTAACTTTGGAAACAGATAGAGAGCAACATCGTTAGTGTCC 791
Db 672 -----AGATAGAGAGCAACATCGTTAGTGTCC 699
Qy 792 TGAATCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGATACAGTTCCAGGGGA 851
Db 700 TGAATCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGATACAGTTCCAGGGGA 759
Qy 852 TTCGCTCGGTGAGTGATCACAATGCTCGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 911
Db 760 TTCGCTCGGTGAGTGATCACAATGCTCGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 819
Qy 912 TGTGTTTCGAACAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTGCGAGCAA 971
Db 820 TGTGTTTCGAACAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTGCGAGCAA 879
Qy 972 CGGCTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGAAGAGTGAAGTGAAGTGA 1031
Db 880 CGGCTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGAAGAGTGAAGTGAAGTGA 939
Qy 1032 TCAGTTTTATATGGAATAGCTGGAAGTGTGAGGACAGCACTAATCGCTCGGGTGGTGT 1091
Db 940 TGAATTTTATATGGAATAGCTGGAAGTGTGAGGACAGCACTAATCGCTCGGGTGGTGT 999
Qy 1092 GCAGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTTAAAGCGACACC 1151
Db 1000 GCAGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTTAAAGCGACACC 1059
Qy 1152 ATCTATCGGGTGAACAGAACACCGTGTCAAGGAGCGTCAAGACGTTGAGCTTCTAGC 1211
Db 1060 ATCTATCGGGTGAACAGAACACCGTGTCAAGGAGCGTCAAGACGTTGAGCTTCTAGC 1119
Qy 1212 AAGAGGGCCCATGACCCATGCTCGCCCTCGAGCTGTTCTGTGTTGGAATCCATGGC 1271
Db 1120 AAGAGGGCCCATGACCCATGCTCGCCCTCGAGCTGTTCTGTGTTGGAATCCATGGC 1179
Qy 1272 CGCGTTGCTCTCATGGACCACTGATGGCGCACTGTTGGCGCACTGAGTGAGATGTTCCGCT 1331
Db 1180 CGCGTTGCTCTCATGGACCACTGATGGCGCACTGTTGGCGCACTGAGTGAGATGTTCCGCT 1239
```

```
Qy 1332 CAATACTGCATTTCAAGAACCACTTGGCTCTTTCTAGCAGAGGACGACACCTGATGA 1391
Db 1240 CAATGCTGCATTTCAAGAACCACTTGGCTCTTTCTAGCAGAGGACGACACCTGATGA 1299
Qy 1392 GCTCGGCCCAAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGTGGCTTGGTTTG 1451
Db 1300 GCTCGGCCCAAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGTGGCTTGGTTTG 1359
Qy 1452 CTTGTCTCTTGCAACCGTAGTTTTTTTTTTTTTTT-TTCCCGCAAGTGTGATCGGATGAAGTG 1510
Db 1360 CTTGTCTCTTGCAACCGTAGTTTTTTTTTTTTTTT-TTCCCGCAAGTGTGATCGGATGAAGTG 1419
Qy 1511 AATAAGGCACTTGTGTTCCCTGTCATTTGTACAGTTTTCATATAATCTAATCTACTTCGA 1570
Db 1420 AATAAGGCACTTGTGTTCCCTGTCATTTGTACAGTTTTCATATAATCTAATCTACTTCGA 1479
Qy 1571 AGATGATGCATTTTATAGATGTGGCTTGTGAAAGACAAAAAAAAGAAAAA 1621
Db 1480 AGATGATGCATTTTATATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1530

RESULT 4
US-10-425-115-138337
; Sequence 138337, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 138337
; LENGTH: 1845
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57640C.1
US-10-425-115-138337
```

```
Query Match 81.3%; Score 1321.8; DB 20; Length 1845;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1574; Conservative 0; Mismatches 17; Indels 212; Gaps 3;

Qy 11 AGCTTCTGCTCTCTCGCGCGCGCGGAGGCAATCATCATCTTCTAGCTCATCAAT 70
Db 43 AGCTTCTGCTCTCTCGCGCGCGCGGAGGCAATCATCATCTTCTAGCTCATCAAT 102
Qy 71 CTATTCCGATGACACCGTGCCTCCAAAGCCACAGCAGGTGGGCACTCAAGGCAAGGCTCG 130
Db 103 CTAATTCGATGACACCGTGCCTCCAAAGCCACAGCAGGTGGGCACTCAAGGCAAGGCTCG 162
Qy 131 CACCCCGGCGATCGGCGCTTGTGAGTTGGCCAGGCTCTCTCCCTCCGCTTCG 190
Db 163 CACCCCGGCGATCGGCGCTTGTGAGTTGGCCAGGCTCTCTCCCTCCGCTTCG 222
Qy 191 CCGTGACCGCTCGCGCACTGCTCGCTAGAGGTGAAGGCACTCTGGAAACACGTTTGGAA 250
Db 223 CCGTGACCGCTCGCGCACTGCTCGCTAGAGGTGAAGGCACTCTGGAAACACGTTTGGAA 282
Qy 251 ACTACTTTCAGGTTGCAACCTATCGTGAATCTCAATGAGGGGTGGTGTGTTGTGTATCA 310
Db 283 ACTACTTTCAGGTTGCAACCTATCGTGAATCTCAATGAGGGGTGGTGTGTTGTGTATCA 342
Qy 311 GTGGTTGTCACCTAGAAATTCACCTCACTGAGGCGAGACCTACAAAGTTGAATCCGATCGAA 370
Db 343 GTGGTTGTCACCTAGAAATTCACCTCACTGAGGCGAGACCTACAAAGTTGAATCCGATCGAA 402
```



Db	410	CCAGAA	TTCTCTCTC	ACTGAGG	CAGCATG	CAAGTAG	AAC	TCGATAG	AAGAC	GTCCGG	GT	469
Qy	383	AGACG	AGATAA	CACTCC	ACAAGG	AGAGACT	GTATAC	ATG	CAAAAT	TTCTGT	CAGG	442
Db	470	AAAGT	AGNAT	TTACAA	CCCCA	AGAAAGG	AGACTGT	ATAC	GCAAA	TTCTAT	CAGG	529
Qy	443	ATGAAG	GGGTGA	CTATCT	GTGGA	ACGCAAT	CTTTGT	TTATG	TGCC	AAAA	CACAGAT	502
Db	530	ATGAT	GGGAT	GACTACT	CTGGT	ACACCAAT	TCACG	CTTTGT	TGCC	AAAA	CACAGAT	589
Qy	503	GCATG	ATCAC	CGTGA	ATAGCC	CAATGTG	TACCG	ACCTTCT	CTATG	CAGAG	CGCACT	562
Db	590	GTGTG	ATTTAC	AGTGA	AAATGT	CTAAG	GGGTAC	AGACCA	TCCTCC	ATG	CGAAT	649
Qy	563	ACTTCA	AGATG	ACGGT	TTAGAG	CTGTAC	AGGGAG	GTGGAG	GTCTCG	GGCAG	AAAAAC	622
Db	650	ACTTCA	AGATG	ATGGA	TTTAGAG	CTGTG	CAGG	GGGTGG	AGGT	CATCAG	CCAGAG	709
Qy	623	TTGGA	AGGTGG	CTGC	AGGG	CCCTCC	CAAG	AAAAATTT	TTAA	AGCTCA	AATGTG	682
Db	710	TTGCG	AGGTGG	CTGC	AGG	AGCTCT	TGCA	AGAAAA	TTCT	AAAG	CTCAAA	769
Qy	683	AGATCT	TGTCG	TTTGT	TTCC	AAAGTG	CAATC	AGGTTGT	CTCC	CAGA	AGACG	742
Db	770	AGATCT	TGGCA	TTTGT	TTCT	AAAGTG	CAC	CAAGTCG	TACTT	CCAG	AGATG	829
Qy	743	ATGGG	CTGTAA	CTTTG	GAA	CAGATAG	AGAG	CAACAT	CGT	PAGAT	CTCTG	802
Db	830	ATGAG	ACTGT	AACTT	CGAA	CATATAG	AGAG	CAACAT	CGT	TAGAT	CTCTG	889
Qy	803	ACG	CAGAG	AGATG	ATAG	CCATC	CA	GAGTAC	CGAG	TTTCG	AGGG	862
Db	890	ATG	CAGAG	AGATG	ATG	CTG	CTCC	CAATG	AT	ACGG	TAG	949
Qy	863	GAGT	ATC	ACAT	TGCG	TCTAG	AAAC	CGTTCT	CCCG	GGCTCG	GTTCCT	922
Db	950	GGG	CGTC	CACATG	CATTG	CAAG	AAATG	TTCT	CTCG	TGGTCT	CTCTG	1009
Qy	923	AGCT	CGAA	TCG	AACTG	GC	AAAAAG	CTATG	CTTTCT	ATCTC	TCG	982
Db	1010	AACT	TGA	AGCTG	AACTG	CA	AAAAAG	CCATG	CTTTCT	CTCTC	G	1069
Qy	983	TTG	CAG	CGGAT	TCG	CGGG	ACCG	ACTTG	AC	GGAAG	TGAG	1042
Db	1070	TTG	CAATG	GGG	TTG	CTGTG	TA	CGACTT	TACT	CG	AAATG	1129
Qy	1043	TGGA	TAA	GGCTG	GGA	GTG	AGG	CA	CGG	ACTAAT	TCG	1102
Db	1130	TGG	ATG	AGG	CTG	GAATG	TG	AGG	CA	CGAA	CTAAT	1189
Qy	1103	TAT	CGA	ATG	TG	GATG	TC	CACTT	CAAA	GTG	CTTTT	1162
Db	1190	TAT	CAAA	TGG	TGA	ATATAT	TTACT	TTCA	AAAG	TGG	CTTTT	1249
Qy	1163	TGAA	ACAG	AA	CCG	GTG	CAAG	GGAG	CGCTC	AGAA	CGTTG	1222
Db	1250	AGA	AGCA	AAAT	TACT	GTG	TC	AA	GGG	CGATG	AGG	1309
Qy	1223	ATG	ACCA	TG	CGT	CG	CCCT	CTG	AG	CTG	TTT	1282
Db	1310	ATG	AC	CCCT	CTG	TTG	TG	CTC	CTG	AG	CTG	1369
Qy	1283	TCAT	G	AC	CA	CG	CTG	ATG	CG	CG	CA	1342
Db	1370	TGAT	TG	GA	CC	CA	CT	ATG	GG	CA	TATG	1429
Qy	1343	TTCA	AG	AA	CC	AG	TTG	GG	CTCT			
Db	1430	TACA	AG	AG	CC	CA	TTG	GG	CTCT			

```

; Sequence 1, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
;
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chromatase Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-677-179-1

```

```
Db 716 ATGAGACTGTAACTTTGGAAACATATAGAGAGCAACATCGTTAGATGCTCTGATCCAGAAT 775
Qy 803 AGCCAGAGAGATGATAGACGCAATCGACAGATACGAGTTTCGAGGGGATTCGGTCGGTG 862
Db 776 ATGACAGAGAAGATGATGTCCTCCATTTGATACGGTACGAGTTAGAGAGAGATCAATTTGGTG 835
Qy 863 GAGTGATCATCGTCGCTAGAAACGTTTCCTCGCGGGCTCGGTCTCTCTGTTTCGACA 922
Db 836 GGGTCGTACATGCAATTCGCAAGAAATGTTCTCTGTTGCTTGGCTCTCTCTGTTTGACA 895
Qy 923 AGCTCAATCCGAATCGGCAAAAGCTATGCTTTCTATTCTTCGAGCAACGSGTTTCGAGA 982
Db 896 AACTTGAAGCTGAATCGGCAAAAGCCATGCTTTCTCTCTGCAAGCAAGGGGTTTGAGA 955
Qy 983 TTGGCAGCGGATTCGCGGGGACCGACTTCACAGGAGTGAGCATATGATGATGTTTATA 1042
Db 956 TTGGCAGTGGGTTGGCTGGTACGGAATTTACTGGAAGTGAGCATATGATGATGTTTATA 1015
Qy 1043 TGDATAAGGCTGGAAGTGTGAGACACGCACTAATCGCTCGGTGGTGTGCAAGGAGGGA 1102
Db 1016 TGGATGAGCTGGAATGTGAGGACACGAACTAATCGCTCAGCGGTGTTCAAGGAGGGA 1075
Qy 1103 TATCGAATGTTGAGATGTGCACTTCAAAAGTGTCTTTAAAGCGCACCAATCTATCGGGG 1162
Db 1076 TATCAAAATGGTGAATTTATTTACTTCAAAAGTGCTTTTAAAGCCACAGCAACTATCGGAA 1135
Qy 1163 TGAACAGAACCCGTGTCAGGAGCGTCAAGAGCTGAGAGCTTCTAGCAAGAGGCGCC 1222
Db 1136 AGAAGCAAAATAGTGTCAAGGGAGCATGAGGATGTTGAACCTTTTGGCAAGGGGCGCC 1195
Qy 1223 ATGACCCATCGCTCGCCCTCGAGCTGTTTCCTGTGGTGAATCCATGGCGGTGGTCC 1282
Db 1196 ATGACCCCTGTGTGCTCTCGAGCTGTTCTTATGTTGAATCCATGGTGGTGGTCC 1255
Qy 1283 TCATGACCAAGCTGATGAGCGGACGTGCTCAGTGGGAGATGTTCCGCTCAATACCTGCAC 1342
Db 1256 TGATGACCAAGCTCATGGGGCATATGCCCAGTGTGAGATGTTTCGCTGAACCTTTGCC 1315
Qy 1343 TTCAAGAACCACTTGGCTCT 1362
Db 1316 TACAAGAGCCCATTTGGCTCT 1335
```

## RESULT 7

```
US-10-425-114-32244
; Sequence 32244, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kvalic, David K.
; APPLICANT: Sreen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32244
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73278F07_FLI
US-10-425-114-32244
```

Query Match 53.3%; Score 867.2; DB 18; Length 1603;  
Best Local Similarity 80.4%; Pred. No. 9.1e-260;  
Matches 1029; Conservative 0; Mismatches 248; Indels 3; Gaps 1;

```
Qy 83 GCACCGTGTCCCAAGCCACAGCAGGTGGCGCACTCACGGGCGACGGCTCGCACCCCGCGGA 142
Db 44 CGCCCGTGTGCGAGCGCGGGTGTCCGCCAGGGCGTCCACAGGTTTCTCCCCCGGGA 103
Qy 143 TCGCGCCTTGTGTGAGTTTCCCGAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 202
Db 104 TAGCGCGCTCTCCCGAGT---CCGCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 160
Qy 203 GCGCACCTGCTCGCTTACAGGTGAGGCGATCTGGAACACGCTTTGGAACACTCTTTTCAGG 262
Db 161 GCGCGCGCGCAGCCCTAGAGGTGAAGGCATCGGGAATGTGTTCGGGAACACTCTTTTCAGG 220
Qy 263 TTGCAACCTATGTTGAACTCTCATGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 322
Db 221 TTGCAACCTATGCGGAATCCCATGGAAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 280
Qy 323 CTAGAATTTCCATCTCATCTGAGGAGACCTTCAAGTTGAACTCGATCGAAGAGCGGCCGAC 382
Db 281 CCAGAAATTCCTCTCTCACTGAGGAGACATGCAAGTAGAACTCGATAGAAGAGCTCCGGGTC 340
Qy 383 AGAGCAGAAATTAACCTCCACAGGAGGAGACTGATACATGCAAAATTTCTGTCCAGGAGAC 442
Db 341 AAAGTAGAATTTACAAACCCCAAGAAAGGAGACTGATACATGCAAAATTTCTATCAGGAGAC 400
Qy 443 ATGAAGGGGTGACTACTGGAACGCCAAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 502
Db 401 ATGATGGGATGACTACTGTGTACCAATTCAGCTTTGTGTCCAAACACAGATCAAGGG 460
Qy 503 GAGTGATCAACGTTGAAATAGCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 562
Db 461 GTGTGTGATTAAGTGAATGTCTAAGCGGTACAGACCATCCATGAGATGCAACCTATG 520
Qy 563 ACTTCAAGTACGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 622
Db 521 ACTTCAAGTATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 580
Qy 623 TTGGAAGGGTGGCTGCGAGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGTGTGT 682
Db 581 TTGCGAGGGTGGCTGCGAGGAGCTCTTGTCAAGAAAAATTTCTAAGCTCAAAATGTGTGT 640
Qy 683 AGATCTGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 742
Db 641 AGATCTTGGCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 700
Qy 743 ATGGGTCTGTAACTTTGGAACAGATAGAGACCAATCGTTAGATGCTCTCTGATCCAGAT 802
Db 701 ATGAGACTGTAACTTTGGAACATATAGAGAGCAACATCGTTAGATGCTCTGATCTAGAAT 760
Qy 803 ACGCAGAGAAGATGATAGACCAATCGACAGATACGAGTTTCGAGGGGATTCGGTCGGTG 862
Db 761 ATGCAAGAGAAGATGATGCTGCCATGTATACGGTACGAGTTAGAGAGATTCATTTGGTG 820
Qy 863 GAGTGATCACAATGCTGTAGAAAACGTTCTCTCGGGGCTCGGTTCCTCTGTGTGTGTGTGTGT 922
Db 821 GGGTCGTACATGCAATGTGCAAGAAATGTCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 880
Qy 923 AGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATCTCTCGAGCAACGSGTTTCGAGA 982
Db 881 AACTTGAAGCTGAATCGCAAAAGCCATGCTTTCTCTCTGCAAGCAAGGGGTTTGAGA 940
Qy 983 TTGCGAGCGGATTCGCGGGGACCGACTTTCAGAGAAAGTGAGCATAATGATGATGATTTTATA 1042
Db 941 TTGCGAGTGGGTTTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1000
Qy 1043 TGGATAAGGCTGGAAGTGTTCAGGACACGGAATTAATCGCTCGGGTGGTGTTCAGGGAGGGA 1102
Db 1001 TGGATGAGGCTGGAATGTGTAGGACACGGAATTAATCGCTCAGGGCGGTGTTCAGGGAGGGA 1060
Qy 1103 TATCGAATGTTGAGATGTTGCACTTCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1162
Db 1061 TATCAAAATGTTGAAATTTATTTACTTCAAAAGTGGCTTTTAAAGCCCAACCAACTATCGGA 1120
Qy 1163 TGAACAGAACACCGTGTCAAGGGGAGCGTCAAGACGTTGAGCTTCTAGCAAGAGGCGCC 1222
```

Db 1121 AGAAGCAAAATACGTGTCAAGGAGCATGAGATGTTGAACCTTTTGCAAGGGGGCGCC 1180  
Qy 1223 ATGACCCATCGGTGCGCCCTCGAGCTGTTCTGTGTGGAATCAATGSCCGGTGGTTC 1282  
Db 1181 ATGACCCCTGTGTGTCCTCGAGCTGTTCTATGTGGAATCCATGGCTCGCTGGTTC 1240  
Qy 1283 TCATGGACCACTGATGCGGACGCTGGCTCAGTGGAGATCTTCGCGCTCAATACTGCAC 1342  
Db 1241 TGATGGACCACTGATGCGGACATATTGCCAGTGTGAGATGTTCCGCTGAACCTTGCCC 1300  
Qy 1343 TTCAAGAACCACTGTTGGCTCT 1362  
Db 1301 TACAAGACCCATTGGCTCT 1320

## RESULT 8

US-10-437-963-99207/c  
; Sequence 99207, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 99207  
; LENGTH: 1353  
; TYPE: DNA  
; ORGANISM: *Oryza sativa*  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_9703C.1  
US-10-437-963-99207

Query Match 49.6%; Score 806.4; DB 19; Length 1353;  
Best Local Similarity 78.4%; Pred. No. 8.8e-241;  
Matches 1010; Conservative 0; Mismatches 236; Indels 42; Gaps 2;  
Qy 75 TCCGATGACGACCGTGCCCAAGCCACAGCAGGTGGCGCACTCAGGGCAGCGCTCGCAC 134  
Db 1285 TCCGATGCGCGCCCAAGCTGCTCGCAGCGGTGGCGCGCTCTCCCGCGGGGGG 1226  
Qy 135 CGCGCGATCGCGCGCTTGTGGAGTTTGGCCAGCTCTCTCCCTCGCGTTCGCGT 194  
Db 1225 CGCGGGTTCCGCGCTTCCGAGTTCGCGCGCT-----TCCTCCGCTTCTCGT 1172  
Qy 195 GCACGCTGCGCACTGCTCGCTAGAGTGAAGCATCTGGAACACAGTTTGGAACTA 254  
Db 1171 CGCGCGCGCGCGCGCTTCCGCTAGAGTGAAGCGTCTGCAATGATTATTGGGA 1112  
Qy 255 CTTTCAGGTTGCAACTATGTGTAATCTCATGGGGGTGGTGTGGTTGTGTTATCAGTGG 314  
Db 1111 CTTTCAGGTTGCAACTATGTGAGAGTCTCATGAGCGGTGTGGTTCGTTAATCAGTGG 1052  
Qy 315 TTGTCACCTAGAAATTCACCTACCTGAGGAGACCTACAAGTTGAATCTGATCGAAGACG 374  
Db 1051 ATGCCACCCCAATCCCACTTACTTGAAGCAGATATGCAAGTAGAACTCGACCGGAGACG 992  
Qy 375 GCCCGGACAGACAGCAATAAATCCACAGGAAGAGAGTGTATCATATCAAAATTCGTCTC 434  
Db 991 GCCAGCCAGACAGCAATAAACACCCCAAGAAAGAGAGTGTACACTTGCATAATTCCTTC 932  
Qy 435 AGGGACACATGAAGGGGTGACTTACTTGGAAACGCAATCTTGTATTGTTCCTCCAAACACAGA 494

Db 931 AGGCACACATGAAGGAATGACCACTGGGACACCAATTCATGTGTTTTCCTCCGAACACAGA 872  
Qy 495 TCAATAGGCAGTGTATCACCCTGAAATAGCAATGTGTACCGACCTTCTCATGCAGACGC 554  
Db 871 TCAGAGAGGGGGTGTATACAGTGAATGGTAAAGCCCTACAGACCTTTCATCATGAGATGC 812  
Qy 555 AACTTATGACTTCAAGTACCGTGTAGAGTGTACAGGGAGGTGGGAGGTCTCTCGGCAG 614  
Db 811 AACTTATGACTTCAATACCGTGTAGAGCAGTGCAGGGAGGTGGAAAGATCATCAGCAAG 752  
Qy 615 AAAAACCGTTTGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAAATCTTAAAGCTCAAATG 674  
Db 751 AGAGACCATTTGAAGGGTGGCTGCAGGAGCTCTTGCAAGAAAAATCTTAAAGCTCAAATC 692  
Qy 675 TGGATTAGAGATCTTGTGCTGTTTGTTCAAAAGTGCATCAGGTTGTCTCCAGNAGACGC 734  
Db 691 TGGAGTAGAGATCTTGGCGTGTGTCCTCAAGGTGCATCAAGTTGTACTACCGAAGATGC 632  
Qy 735 GGTGTGATTATGGGTCTGTAACTTTGGAAACAGATAGAGCAACATCGTTAGATGCTCTGA 794  
Db 631 CGTTGATTATGACACTGTAACTTGAACAGATAGAAAGCAACATTTGTTAGATGCTCTGA 572  
Qy 795 TCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGATGACAGTTCGAGGGGATTC 854  
Db 571 TCCAGAAATATGCACAGAAAGATGATTGATGCAATCGATAAAAGTACGAGTTAGAGGTATTC 512  
Qy 855 GGTGCGTGGAGTGTATCATGCTCGCTAGAAAGTTCCTCGCGGGCTCGGTTCTCTGT 914  
Db 511 GATTGGTGGTGGTGTGATGATGCAATGCAAGAAATGTTCTCGTGGGATTTGGCTCTCTGT 452  
Qy 915 GTTCGACAAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCTCGCAACACGG 974  
Db 451 ATTTGACAAACTTGAGCTGAAATTTGGCGAAAGCTATGCTTTCTCTCTCGCAACAGGG 392  
Qy 975 GTTCGAGATTGGCAGCGGATTCGCGGGACCGAATTGACAGGAAGTGAAGCATAATGATGA 1034  
Db 391 GTTTGAGATCGGCAGTGGAATTTGCAGGTACTGACTACACTGGAAGTGAAGCATAATGATGA 332  
Qy 1035 GTTTTATATGATGAAGCTTGAAGTGTGAGACACGAGCTAATCGCTCGGGTGTGTGCA 1094  
Db 331 GTTCTATATGATGAGGCTGGA-----A 308  
Qy 1095 GGGAGGATATCGAATGTTGAGATTGTGCATTTCAAAAGTTGCTTTTAAAGCCGACCATC 1154  
Db 307 TGGAGGATATCAATGGTGAATATATATCTTCAAGTAGCTTTCAAGCCACCGGAC 248  
Qy 1155 TATCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTTGAGCTTTCTAGCAAG 1214  
Db 247 TATTGGGAAGAAACACATACTGTTTCAAGGGAGCATGAGGATGTTGAACTTTTAGCAAG 188  
Qy 1215 AGGGCGCATGACCCATGCGTCCGCCCTCGAGCTGTTCTGTGTGGTGAATCCATGGCGGC 1274  
Db 187 GGGCGCCACAGCCCATGTTGTCTCGCTGTGCCGATGGTGGAGTCCATGGCGGC 128  
Qy 1275 GTTGGTCTCATGACCAAGCTGATGGCGCACTGGCTCAGTGGCAGATGTTTCGGCTCAA 1334  
Db 127 ATTAGTCTCATGACCAAGCTGATGGCGCATTTGCTCAATGTGAGATGTTTCCACTGAA 68  
Qy 1335 TACTGCACTTCAAGAACCAAGTTGGCTCT 1362  
Db 67 CTTGCTCTACAGAACCAAGTTGGCTCT 40

## RESULT 9

US-10-677-179-11  
; Sequence 11, Application US/10677179  
; Publication No. US20040082050A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Choriastate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A



;; CURRENT APPLICATION NUMBER: US/10/677,179  
;; CURRENT FILING DATE: 2003-10-02  
;; PRIOR APPLICATION NUMBER: US/09/743,207  
;; PRIOR FILING DATE: 2001-01-04  
;; PRIOR APPLICATION NUMBER: PCT/US99/16353  
;; PRIOR FILING DATE: 1999-07-20  
;; PRIOR APPLICATION NUMBER: 60/093,611  
;; PRIOR FILING DATE: 1998-07-21  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: Microsoft Office 97  
;; SEQ ID NO 11  
;; LENGTH: 966  
;; TYPE: DNA  
;; ORGANISM: Oryza sativa  
US-10-677-179-11

Query Match 38.2%; Score 621; DB 18; Length 966;  
Best Local Similarity 80.4%; Pred. No. 7.4e-183;  
Matches 742; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

QY 75 TCCGATGACGACCGTCCCAAGCCACAGCAGGTGGCGCACTCACGGGACGGCTCGCACC 134  
DB 41 TCGATGGCGCGCCCAACGTGTGCGAGCGGTGGCGCGTCTCCCGCGGGCGCG 100  
QY 135 CGCGCGATCGCGCGCTTGTGTGAGTTTGCCTCCAGCTCTCTCTCCCTCCGCTTCCCGGT 194  
DB 101 CGCGGGTTCGCGCTTCCCGAGTCCGCGCGCT-----TCCTCCGCTTCTCGT 154  
QY 195 GCACCGCTCGCGACGTGCTGCTAGAGTGAAGCACTTGAAACACCTTTTGGAACTA 254  
DB 155 CGCGCGCGCGCGCGCTCGCTAGAGTGAAGCGTCTGCAATGTATTTGGGACTA 214  
QY 255 CTTTCAGGTGCAACTTGTGATCTCATGGGGTGTGTGTGTGTATCATGTTG 314  
DB 215 CTTTCAGGTGCAACTTGTGAGAGTCTCATGAGCGGTGTGTGTGTGTATCATGTTG 274  
QY 315 TTGTCCACCTAGAAATTCCTCACTGAGCGAGACCTTACAAAGTTGAATCGATCGAAGACG 374  
DB 275 ATGCCACCCAGAAATCCCACTTACTGAGCAGATATGCAAGTAGAACTCGACGGAGACG 334  
QY 375 GCCTGGACAGAGAGAAATTAACCTTCAAGAGGAGACTGATACATGCAAAATTTCTGT 434  
DB 335 GCAGCGCCAGAGCAGAAATTAACCAACCCCAAGAAAGAGACTGACACTTGCAAAATTTCTTC 394  
QY 435 AGGACACATGAAGGGTCACTACTGGAAGCGCAATTTCTGTATTTGTTCCCAACACAGA 494  
DB 395 AGGACACATGAAGGAATGACCACTGGGACACCAATTCATGTTTGTCCGGAACACAGA 454  
QY 495 TCAATATAGGACGTGATCAACGTAATAGCAATGTGTACCGACCTTCTCATGCAAGACG 554  
DB 455 TCAGAGAGGGGTGATTAAGTGAATGCTAAGGCTTACAGCTTACATGCAAGTGC 514  
QY 555 AACTTATGATCTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGAGGTCTTCGGGAG 614  
DB 515 AACTTATGATCTTCAATACCGTGTAGAGCTGTACAGGAGGTGGAGATCATCAGCAAG 574  
QY 615 AAAAACCCTTGAAGGGTGTGCGAGGGGCTTCCCAAGAAATTTCTTAAGCTCAATG 674  
DB 575 AGAGACCAATTTGAAGGGTGTGCGAGGAGCTCTTGCAAAAGAAATTTCTTAAGCTCAATC 634  
QY 675 TCGATTAGAGATCTTGTCTGTTTGTTCCTCAAGTGTATCAGTGTGTCTCCCAAGAGACG 734  
DB 635 TCGAGTAGAGATCTTGTGCTTGTGCTCAAGTGTATCAAGTTGTACTACCAAGAGATGC 694  
QY 735 GGTTCATTATGAGTCTGTAACTTTTGAACAGATAGAGCAACATCGTTTGTAGATGCTTGA 794  
DB 695 CGTTGATTATGACACTGTAACTTGAACAGATAGAGCAACATCGTTTGTAGATGCTTGA 754  
QY 795 TCCAGATACGAGAGAGATGATAGACCAATTCAGACAGTACAGTTTCAGGGGATTC 854  
DB 755 TCCAGAAATATGACAGAGAGATGATGATGCACTTCGATAAAGTACAGTTTGTAGAGTTC 814  
QY 855 GGTTCGGTGGAGTATCAATGCGTTCGCTAGAAACGTTCTCTCGCGGCTCGGTTCTCTCTGT 914

DB 815 GATTGTTGTTGGTGCATGCAATGCAAGAAATGTTCTCTGCGGATTTGGCTCTCTGT 874  
QY 915 GTTCCAGAGCTCGAATCGGAATCGGAAGCTATGCTTTCTATTTCTGCGAGCAACGG 974  
DB 875 ATTTGACAAACTTTGAGGCTGAATGCGGAAGCTATGCTTTCTCTTCTGCAAGCAAGGG 934  
QY 975 GTTCCAGATTGCGCAGCGGATTTCG 997  
DB 935 GTTTCAGATCGCAGTGGATTTG 957

## RESULT 10

US-10-425-114-8982

;; Sequence 8982, Application US/10425114

;; Publication No. US20040034888A1

;; GENERAL INFORMATION:

;; APPLICANT: Liu, Jingdong

;; APPLICANT: Zhou, Yihua

;; APPLICANT: Kovalic, David K.

;; APPLICANT: Screen, Steven E.

;; APPLICANT: Tabaska, Jack E.

;; APPLICANT: Cao, Yongwei

;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

;; FILE REFERENCE: 38-21(53313)B

;; CURRENT APPLICATION NUMBER: US/10/425,114

;; CURRENT FILING DATE: 2003-04-28

;; NUMBER OF SEQ ID NOS: 73128

;; SEQ ID NO 8982

;; LENGTH: 1674

;; TYPE: DNA

;; ORGANISM: Glycine max

;; FEATURE:

;; OTHER INFORMATION: Clone ID: 700829731\_FLI

US-10-425-114-8982

## Query Match

35.5%; Score 578; DB 18; Length 1674;

Best Local Similarity 67.8%; Pred. No. 2.9e-169;

Matches 809; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 164 CCCAGCCTCTCTCTCCCTCCCTCGCTTCCGCGTGCACCGTGCACCTGCTCGCTTAGAGG 223  
DB 90 CCGTCCCGCGCTACCTCGGCTCTCACTCCGCTCTCCGCTTCCCAAGAGACTTCACA 149  
QY 224 TGAGGCACTGGAACACACGTTTGGAACTACTTTTCAGGTTGCAACCTATGCTGATCTC 283  
DB 150 TACAGCGCGCTGGGAGTACCTATGGAATCACTTTTCGTTTACACATATGGGAATCAC 209  
QY 284 ATGGGGTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 343  
DB 210 ATGGAGGAGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 269  
QY 344 CAGACCTACAAAGTTGAACTCCGATCGAAGACCGCGGACAGCAGACAGATAACCTCCACA 403  
DB 270 CTGATATGCAATGGAATCTTCAAGAGGAGCGGAGGTCAGAGCCGATTAACACTCTTA 329  
QY 404 GGAAGGAGCTGATACATGCAAAATTTCTGTGAGGACACATGAAAGGGGTGACTACTGAA 463  
DB 330 GAAAGGAGCTGATACATGTAATAATTTTCAGGAGTTTCCGAAGGAATCACTACTGAA 389  
QY 464 CGCAATTTCTGTTTATTTGTTCCCAACACAGATCAATAGGCGAGTATCACCGTGAATAG 523  
DB 390 CTCGAATTCATGATATCTGTATCCCAATCTGATCAAGAGGAGCATGACTATAGCGAGATGG 449  
QY 524 CCAATGTGTACCGGACCTTCTCATGACAGCGCAACTTATGACTTCAAGTACCGTGTAGAG 583  
DB 450 CAGTAGCTTATAGGCCCTCCCATGCAATGCTACTATGATGATGATGATGATGATGATG 509  
QY 584 CTGTACAGGAGGTGGAGGTCTCTCGGCAAGAAACCGTTTGGAGGGGTGCTGCAAGGG 643  
DB 510 CAGTTCCAGGTGGTGGTAGATCTTCTGCAAGAGAAACAAATTTGGAAGGGTGTCTCTGTTG 569







```
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-677-179-5

Query Match      30.8%; Score 500.6; DB 18; Length 1015;
Best Local Similarity 80.3%; Pred. No. 3.6e-145;
Matches 587; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 631 GTGGCTGAGGGGCCCTCCCAAGAAAATTTCTTAAGCTCAAAATGGAATTAGAGATCTTG 690
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 5 GAGGCTGAGGAGCTGTGTCGAAGAAAATTTCTTAAGCTCAAAATGGAATTAGAGATCTTA 64
QY 691 TCGTTTGTTCACAAAGTCATCAGCTGTGCTCCAGAAAGCGGTTGATTATGGGTCT 750
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
G5 65 GATTGTGTTCCAAAGTGATCAAGTGGTACTTCTGAAGACGCGATTGATTATGAACCT 124
QY 751 GTAACTTTGGAACAGATAGAGCAACATCTGTAGATGCTCTGTATCCAGAGTACGCAGAG 810
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 125 CTTACCTGGATCAGATAGAGCAACATTTGTAGATGCTCTGTATCCAGATATATGCACAG 184
QY 811 AAGATGATAGAGCAATCGACAGATACGAGTTCGAGGGGATTCGGTTCGGTGGAGTGATC 870
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 185 AAGATGATGATGCAATGATAAAGTACGAGTTAATGGGAATTCGATTGGTGGGGTGGTC 244
QY 871 ACATCGCTCGCTAGAAAGTTCCTCGCGGGCTCGGTTCTCTGTGTTTCGACAGCTCGAA 930
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 245 ACATGCAATGGCCAGAAATGTTCTCGTGGGCTTGGCTCTCTGTATTTGACAAACTTGA 304
QY 931 TCCGAATCGCAAAAGCTATGTTTCTATTTCTCGGAGCAACGGGTTTCGAGATTGGCAGC 990
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 305 GCTTACTGGCAAGGCTATGCTTCTCTCTCGAAGCAAGGGGTTTGAGATCGGTAGT 364
QY 991 GGAATCGCGGAGCCGACTTTGACAGGAAGTGAGCAATAATGATGATTTTATATGATGAAG 1050
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 365 GGAATTCGAGGTACTGACCTAACTGGAAGTGAGCATACGATGAGTTCTATATGACGAG 424
QY 1051 GCTGGAAGTTCAGACACGAGCTAATCGCTCGGGTGGTGTGCGAGGAGGGATATCGAAT 1110
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 425 GCTGGAATGTAGAAACAGAACCAATCGCTCGGGCGGTGTACAGGAGGGATATCAAA 484
QY 1111 GTTGAGATTGTCACTTCAAGTTGCTTTTAAAGCCGACACCATCTATCGGGGTGAAACAG 1170
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 485 GGTGAACCTATATACTTCAAGTAGCTTTCAAGCCACAGCAACTATTTGGGGAAGCA 544
QY 1171 AACACCGTGTGAAGGAGGTGAGAACTGTGAGCTTCTAGCAAGAGGGCGGCATGACCCA 1230
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 545 AATACTGTAAACAAGGGATCATGAGGATATCGAACTTCTGACAAAGGGGTGCGCATGACCCA 604
QY 1231 TGCTCGCCCTCGAGCTGTTCTCTGTGGATCCATCGCGCGCTTGGTCTCATGGAC 1290
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 605 TGTGTGCTCCCTCGGGCTGTTCCAAATGGTGGAGAGATGGCTGCAATGGTCTCATGGAC 664
QY 1291 CAGCTGATGGCGACGTGGCTCAGTGCAGAGATGTTTCGGGCTCAATPATCTGCACCTCAAGAA 1350
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 665 CAGCTGATGGCACATGTGCTCAGTGCAGAGATGTTCCGCTGAACTTCGCTTACCAAGAA 724
QY 1351 CCAATTGGGCTC 1361
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 725 CCAATCGGCTC 735

RESULT 14
US-10-425-115-171487
; Sequence 171487, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 171487
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(724)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_87979C.1
US-10-425-115-171487
```

```
Query Match      26.2%; Score 425.6; DB 20; Length 724;
Best Local Similarity 79.1%; Pred. No. 8.6e-122;
Matches 542; Conservative 0; Mismatches 137; Indels 6; Gaps 3;

QY 168 AGCCTCCTCCTCCTCGCTTCGCGTGGCAGCGCTGCGGCACCTGCTCGCTAGAGTGAA 227
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 41 ACCCTCAACCAACGAGGCAACCAAAACCTTCTCCGTCGCGCGCTCCAGCCTTAGAGTGAA 100
QY 228 GGCATCTGGAACACACGTTTTCGAAACTACTTTTCAGGTTGCAACCTATGCTGAATCTCATGG 287
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 101 GGCATCAGGAATATGTGTTTCGGAACTACTTTCAGGTTGCAACCTATGCGAATCCCATGG 160
QY 288 GGGTGGTGTGGTTGTATCAGTGGTTGTTCACCTTAGAATTCACCTCAGTGGCAGA 347
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 161 AGGGGTGTGGTTGGTTGCTTATCAGTGGCTGCCACCCAGAAATTCCTCTCAGTGGCAGA 220
QY 348 CCTACAAGTTGAACCTGATCGAAGACGCGCGGACAGACAGATACCTCCACAAAGAA 407
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 221 CATGCAAGTAGAACTCGATAGAGACGTCGCGGTCAAAAGTAGAATTAACAACCCCAAGAA 280
QY 408 GGAGACTGATACATGCAAAATTTCTGTCAGGACACATGAAGGGGTGACTACTGGAACGCC 467
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 281 GGAGACTGATACATGCAAAATTTCTATCAGGACACATGATGGGATGACTACTGCTACACC 340
QY 468 AATTCTTTGTTATTTGTCCTCAACACAGATCAAAATAGGACGATGATCACCGTGAATAGCAA 527
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 341 AATTCAAGTCTTTGTGCCAAACACAGATCAAAAGGGGTGGTGATTACAGTGAATGTCTAA 400
QY 528 TGTGTACCGACCTTCTCATGACAGACCAACTTATGACTCAAGTCAAGTGTGTAGAGCTGT 587
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 401 GGGTACAGACCAATCCCATGAGATGCAACTTATGACTTTTAAAGTATGGAGTTAGAGCTGT 460
QY 588 AC-AGGAGGTGGGAGGTCTCTCGGGCAGAAAAACCCGTTGGAAGGGTGGCTGCGAGGGGCC 646
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 461 GCGNGNAGGTGGNAGGTCAATCAGCAGAGAAACCATTGGCAGGGTGGCTGCGAGAGCTC 520
QY 647 TCCCAAGAAAAATTTCTTAAGCTCAAAAT- - -GTGGATTAGAGATCTTTGTGTTGTTTCC 702
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 521 TTGCAAGAAAAATTTCTTAAGCTCAAAATCANNGGAGTGGAGATCNTTGGCATTTGGTTTCT 580
QY 703 AAGTGCATCAGTGTGTGCTCCAGAGACGCGGTTCATTATGCGTCTGTAACCTTTGGAA 762
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 581 AAGTGCACCAAGTCTGACTTCCAGAAAGATGCACTTGAATTATGAGACTGTAAACCTTGAA 640
QY 763 CAGATAGAGAGCAACATCGTTAGATGTCCAGATCCAGATACGACAGAGAGATGATAGAC 822
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 641 CATATAGAGACCAATCGTTAGATGTCTGTATGATGCTCTGATCCA- AATATGACAGAGAGTGTGCT 699
QY 823 GCATTCGACAGAGTACGAGTTCCGAG 847
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D5 700 GCCATTGATACGGTACGAGTTAGAG 724
```

```
RESULT 15
US-10-425-114-13775
; Sequence 13775, Application US/10425114
; Publication No. US2004003488A1
```

GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 13775  
LENGTH: 440  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB143-041-F6\_FLI  
US-10-425-114-13775

Query Match 25.5%; Score 414.2; DB 18; Length 440;  
Best Local Similarity 97.9%; Pred. No. 2.4e-118;  
Matches 430; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
QY 1103 TATCGAATGTTGAGATTGTCACCTTCAAAGTTGCTTTTAAGCGCAGACCATCTATCGGGG 1162  
DB 2 TATCGAATGTTGAGATTGTCACCTTCAAAGTTGCTTTTAAGCGCAGACCATCTATCGGGG 61  
QY 1163 TGAACAGAACCCGTGTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAGGCGCC 1222  
DB 62 TGAACAGAACCACTGTGTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAGGCGCC 121  
QY 1223 ATGACCCATGCGTCGCGCCCTCGAGCTGTTCTGTGTGTGAATCCATGCGCGGTTGGTCC 1282  
DB 122 ATGACCCATGCGTCGCGCCCTCGAGCTGTTCTGTGTGTGAATCCATGCGCGGTTGGTCC 181  
QY 1283 TCATGGACCAAGTCATGAGCGCAGTGGCTCAGTGCAGATGTTGCGGCTCAATCTGCAC 1342  
DB 182 TCGTGGACCAAGTCATGAGCGCAGTGGCTCAGTGCAGATGTTGCGGCTCAATCTGCAC 241  
QY 1343 TTCAAGAACCAAGTTGGCTCTTTCTAGCAGAGGAGCAGACACCTGATGAGCTCGCGCCAA 1402  
DB 242 TTCAAGAACCAAGTTGGCTCTTTCTAGCAGAGGAGCAGACACCTGATGAGCTCGCGCCAA 301  
QY 1403 ATTTATCATTTATCATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTGCTTCTCTTG 1462  
DB 302 ATTTATCATTTATCATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTGCTTCTCTTG 361  
QY 1463 CAACGAGTTTGTGTTTTTTT-TTCCCGCAAGTGTGATGCGATGAAGTAAGGCACT 1521  
DB 362 CAACGAGTTTGTGTTTTTTT-TTCCCGCAAGTGTGATGCGATGAAGTAAGGCACT 421  
QY 1522 TGGTTTCCTGTCATTGT 1540  
DB 422 TGGTTTCCTGTCATTGT 440

Search completed: August 27, 2005, 04:44:23  
Job time : 1026 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: August 25, 2005, 19:19:15 ; Search time 74 Seconds  
(without alignments)  
2242.165 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	429	3	AAY44893
2	1797.5	81.7	440	3	AAY44890
3	1541	70.0	431	8	ADK60818
4	1461	66.4	435	3	AG24018
5	1461	66.4	478	3	AG24017
6	1454	66.1	436	8	ADN73863
7	1454	66.1	436	8	ADT91427
8	1438.5	65.4	435	5	ABB91288
9	1257	57.1	307	3	AAY44895
10	1191.5	54.2	362	8	ADK60817
11	1191.5	54.2	362	8	ADN20024
12	1186.5	53.9	362	8	ADN29835
13	1182.5	53.8	362	8	ADN41780
14	1137	51.7	313	3	AG24019
15	1066	48.5	257	3	AAY44892
16	1035	47.0	360	8	ADS42623
17	981	44.6	356	8	ADN21386
18	938	42.6	376	8	ADK60821
19	938	42.6	376	8	ADN43492
20	934	42.5	378	8	ADP99010
21	930	42.3	432	8	ADK60819
22	929.5	42.2	412	8	ADR86076
23	885.5	40.4	358	6	ABU20907
24	886.5	40.3	473	8	ADN21319
25	862	39.2	361	6	ABU50323

26	860	39.1	361	4	AAU34639	Aau34639 E. coli c
27	860	39.1	361	6	ABU28699	Abu28699 Protein e
28	860	39.1	361	8	ADS45053	Ads45053 Bacterial
29	859	39.0	377	7	ABO66029	AbO66029 Klebsiell
30	858.5	39.0	357	8	ADO57720	AdO57720 Actinobac
31	858	39.0	352	6	ABU33103	Abu33103 Protein e
32	857	39.0	361	4	AAU36115	Aau36115 Klebsiell
33	855	38.9	361	6	ABU31873	Abu31873 Protein e
34	850	38.6	361	6	ABU41141	Abu41141 Protein e
35	850	38.6	362	7	ADF04007	AdF04007 Bacterial
36	846.5	38.5	357	6	ABU39054	Abu39054 Protein e
37	843	38.3	360	8	ADS42725	Ads42725 Bacterial
38	840.5	38.2	357	4	AAU35388	Aau35388 Haemophil
39	840.5	38.2	357	4	AAU35388	Aau35388 Haemophil
40	840.5	38.2	357	5	AAU31486	Aau31486 Haemophil
41	840.5	38.2	357	6	ABU30175	Abu30175 Protein e
42	840.5	38.2	357	8	ADK60820	AdK60820 Neurospor
43	840.5	38.2	359	8	ADS42183	Ads42183 Bacterial
44	840	38.2	361	6	ABU49470	Abu49470 Protein e
45	838	38.1	362	6	ABM70731	Abm70731 Photorhab

## ALIGNMENTS

### RESULT 1

AAV44893  
ID AAY44893 standard; protein; 429 AA.

XX AAY44893;

DT 18-MAY-2000 (first entry)

DE Corn chorismate synthase encoded by clone csiln.pk0050.dil.

XX Corn cDNA clone csiln.pk0050.dil; chorismate synthase;

KW immunological screening; herbicide resistance; antibody; gene mapping;  
KW corn.

OS Zea mays.

XX WO200005353-A2.

PN 03-FEB-2000.

PD 20-JUL-1999; 99WO-US016353.

XX 21-JUL-1998; 98US-0093611P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Cahoon RE, Falco SC, Pember SO;

XX WPI; 2000-182687/16.

XX N-PSDB; AAZ50438.

XX New chorismate synthase polypeptides used to alter the level of the enzyme and thus the level of aromatic to non aromatic amino acids in transformed plants.

XX Claim 6; Page 33-34; 39pp; English.

XX The present sequence is corn chorismate synthase encoded by cDNA clone designated csiln.pk0050.dil. Manipulating either the amount or activity of this enzyme would affect the ratio of aromatic to non-aromatic amino acids in plants. Chimeric gene comprising the cDNA operably linked to regulatory sequences is used to transform host cells to alter the level of expression of chorismate synthase. The gene and its products may be used for immunological screening of cDNA expression libraries and to create transgenic plants which may also be herbicide resistant. Synthetic peptides derived from the gene are also raised antibodies, and used in screening assays to identify inhibitors which may be useful as herbicides

```
SQ Sequence 429 AA;
Query Match 100.0%; Score 2200; DB 3; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-206;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTTVPKQVQVHRSRLAPRAIGALLEFAPAPASSLSRPAVHRCRTARLEVKASGNTFGNYF 60
DB 1 MTTVPKQVQVHRSRLAPRAIGALLEFAPAPASSLSRPAVHRCRTARLEVKASGNTFGNYF 60
QY 61 QVATYGESHGCGVGVISGCPPIPLTEADLQVELDRRPPQOSRITSTRKETDCKILSG 120
DB 61 QVATYGESHGCGVGVISGCPPIPLTEADLQVELDRRPPQOSRITSTRKETDCKILSG 120
QY 121 THEGVTTCPTILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRK 180
DB 121 THEGVTTCPTILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIESNIVRCPPD 240
DB 181 TVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIESNIVRCPPD 240
QY 241 EYAEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGF 300
DB 241 EYAEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSI 360
DB 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSI 360
QY 361 GVKQNTVSREKQNVELLARGRHDPCVAPRVVSVESMAALVMDQLMAHVAQCEMPALNT 420
DB 361 GVKQNTVSREKQNVELLARGRHDPCVAPRVVSVESMAALVMDQLMAHVAQCEMPALNT 420
QY 421 ALQEPVGSF 429
DB 421 ALQEPVGSF 429

RESULT 2
AY44890
ID AY44890 standard; protein; 440 AA.
AC AY44890;
XX
XX
DT 18-MAY-2000 (first entry)
DE
DE
DE
KW Corn chorismate synthase encoded by clone Chpc24.pk0002.hl.fis.
KW Corn cDNA clone chpc24.pk0002.hl.fis; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW corn.
OS
OS
XX
XX
XX WO200005353-A2.
XX
PD 03-FEB-2000.
XX
XX
XX 20-JUL-1999; 99WO-US016353.
XX
XX 21-JUL-1998; 98US-0093611P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon RE, Falco SC, Pember SO;
XX
XX WPI; 2000-182687/16.
XX
XX N-FSDB; AAZ50435.
XX
XX
XX New chorismate synthase polypeptides used to alter the level of the
XX enzyme and thus the level of aromatic to non aromatic amino acids in
XX transformed plants.
XX
```

```
XX
PS
XX
XX Claim 6; Page 29-30; 39pp; English.
CC The present sequence is corn chorismate synthase encoded by cDNA clone
CC designated chpc24.pk0002.hl.fis. Manipulating either the amount or
CC activity of this enzyme would affect the ratio of aromatic to non-
CC aromatic amino acids in plants. Chimeric gene comprising the cDNA
CC operably linked to regulatory sequences is used to transform host cells
CC to alter the level of expression of chorismate synthase. The gene and its
CC products may be used for immunological screening of cDNA expression
CC libraries and to create transgenic plants which may also be herbicide
CC resistant. Synthetic peptides derived from the gene are to raise
CC antibodies, and used in screening assays to identify inhibitors which may
CC be useful as herbicides
XX
SQ Sequence 440 AA;
Query Match 81.7%; Score 1797.5; DB 3; Length 440;
Best Local Similarity 81.9%; Pred. No. 9e-167;
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;
QY 4 VPKPQVQVHRSRLAPRAIGALLEFAPAPASSLSRPAV-HRCRTARLEVKASGNTFGNYF 62
DB 5 VSQPPVVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRRASSLEVKASGNTFGNYF 62
QY 63 ATYGESHGCGVGVISGCPPIPLTEADLQVELDRRPPQOSRITSTRKETDCKILSGTH 122
DB 63 ATYGESHGCGVGVISGCPPIPLTEADLQVELDRRPPQOSRITSTRKETDCKILSGTH 122
QY 123 EGVTTGPTILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRKT 182
DB 123 DGMVTGPTIHVFPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSARETI 182
QY 183 GRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIESNIVRCPPDEY 242
DB 183 GRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIESNIVRCPPDEY 242
QY 243 AEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEI 302
DB 243 AEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEI 302
QY 303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSIGV 362
DB 303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSIGV 362
QY 363 KQNTVSREKQNVELLARGRHDPCVAPRVVSVESMAALVMDQLMAHVAQCEMPALNTAL 422
DB 363 KQNTVSREKQNVELLARGRHDPCVAPRVVSVESMAALVMDQLMAHVAQCEMPALNTAL 422
QY 423 QEPVGS 428
DB 423 QEPVGS 428

RESULT 3
ADK60818
ID ADK60818 standard; protein; 431 AA.
XX
XX AC ADK60818;
XX
XX 20-MAY-2004 (first entry)
XX
XX Saccharomyces cerevisiae chorismate synthase.
DE
XX
XX Apicomplexan parasite; malaria; toxoplasmosis; cryptosporidiosis;
XX eimeriosis; veterinary disease; chorismate synthase.
XX
XX Saccharomyces cerevisiae.
XX
XX US6699654-B1.
XX
XX 02-MAR-2004.
XX
```





PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142309P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145143P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147418P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.

PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.4%; Score 1461; DB 3; Length 435;

Best Local Similarity 67.5%; Pred. No. 8.1e-134;

Matches 276; Conservative 59; Mismatches 62; Indels 12; Gaps 2;

QY 30 PASSSLRFAVHR-----CRT---ARLEVKASGNTFGNYFOVATYGESHGGVGCVI 77

DB 16 PGSSSLFSELRLSSPAVQISLRTQTRKNFQIQATGSSYGTFRVSTFGSHGGVGCII 75

RESULT 5	
AAG24017	
ID	AAG24017 standard; protein; 478 AA.
XX	
AC	AAG24017;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 27534.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125768P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.











PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX  
XX Claim 1; SEQ ID NO 2677; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 362 AA;

Query Match 54.2%; Score 1191.5; DB 8; Length 362;  
Best Local Similarity 62.2%; Pred. No. 1.5e-107;  
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGGVCGVILSGCPRIPLTEADLOVELDRRRPGQSRTSTKET 112  
DB 2 GNTFGSLFRITTFGESHGGGVIIIDGCPPLRLISPEEIQVLDLRRRPGQSKITTPRKEA 61  
QY 113 DTCKILSGTHEGVTGTPIILVNPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVOG 172  
DB 62 DQCEILSGVFEKGTGTPIILVNRKDSQDYNEMAVKYRPSHADATYEAQYGRNWQ 121  
QY 173 GRRSGRTYGRVAAGALPKKILKUCGLEILSFVSKVHQVLPEDAVDYGSVTLQIES 232  
DB 122 GGRSARETIGRVAAGAKKILAQFNGEIVAYVYKSIQDI---EATVDSNTVTLEQVES 178  
QY 233 NVIRCPDEYAEKMDAIDRVVRGDSVGVITCVARNVPLGSPVFDKLESLAKAML 292  
DB 179 NVIRCPDECAKMTIERIDQVLRQKDSIGGVVECAIRNAPKLGSPVFDKLEADLAKAM 238  
QY 293 SIPASNGEISGFAGTDLTGTSEHNDYFMDKAGSVRTNRSGVQGGISNVEIVHFKV 352  
DB 239 SLPATKGEFGSGFAGTLLTSQHNDEYLLDEAGEWRTNRSGVQGGISNGEPIINRI 298  
QY 353 AFKPPISGVKQNTYRERQNVELLARHDPFCVAPRAPVVPVNESMAALVLMQDMAHVAQ 412  
DB 299 AFKPTATIGEQKTVSNIGETTLAAGKRHPDVCVLPRAVPVPMVMAALVLDHLLRFOAQ 358  
QY 413 CE 414  
DB 359 CK 360

RESULT 12  
ADS29835  
ID ADS29835 standard; protein; 362 AA.  
XX  
AC ADS29835;  
XX  
DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #18868.  
DE  
XX Recombinant DNA construct; transformed plant; improved plant property;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
XX bacterial polypeptide.  
OS Bacteria.  
XX US2003233675-A1.  
XX 18-DEC-2003.  
XX 20-FEB-2003; 2003US-00369493.  
XX 21-FEB-2002; 2002US-0360039P.  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
PI WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
FT for expression of a polynucleotide encoding a polypeptide from a  
FT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 18868; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 362 AA;

Query Match 53.9%; Score 1186.5; DB 8; Length 362;  
Best Local Similarity 60.6%; Pred. No. 4.8e-107;  
Matches 220; Conservative 66; Mismatches 74; Indels 3; Gaps 1;  
QY 53 GNTFGNYFOVATYGESHGGVCGVILSGCPRIPLTEADLOVELDRRRPGQSRTSTKET 112  
DB 2 GNTFGSLFRITTFGESHGGGVIIIDGCPPLRLISPEEIQVLDLRRRPGQSKITTPRKEA 61  
QY 113 DTCKILSGTHEGVTGTPIILVNPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVOG 172

Db 62 DTCEILSGVYEGKTLGTPISILVRNKDTRPDQYDEMAQKVRPSHADATYDAKYGRNWQ 121  
 QY 173 GRRSGRKTGVRVAAGALPKKILKLCGLLEILSFVSKVHVVLPEADYGVSTLEQIES 232  
 Db 122 GGRSSARETIGRVAAGAIAKKILRQVANVEIYVKRIKDL---EGVDPNTVTLDOVES 178  
 QY 233 NIVRCPDPEYAEKMDAIDRVVRGDSVGVITCVARNVPRGLGSPVFDKLESELAKAML 292  
 Db 179 NIVRCPDGLADRMIELBQTCGQDSIGGVVCEVARNVPGKLGEPVFDKLEADIADAKVM 238  
 QY 293 SIPASNGFEIGSGFAGTDLTGSHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 352  
 Db 239 SLPASRGFEIGSGFAGTLLTGFEHNDYIDENGEITRTVNRSGGICQGIANGENIILRV 298  
 QY 353 AFKPTPSIGVKONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQDMAHVAQ 412  
 Db 299 AFKPTATIRKEQKTVTRGEETLLAAKGRHDPVLPRAVPVMEAMVALVLCDDLRLHHQ 358  
 QY 413 CEM 415  
 Db 359 CKV 361

## RESULT 13

ADS41780  
 ID ADS41780 standard; protein; 362 AA.

XX ADS41780;

AC ADS41780;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #20210.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAO/Y) CAO 'Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 20210; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or by  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 362 AA;

Query Match 53.8%; Score 1182.5; DB 8; Length 362;

Best Local Similarity 60.6%; Pred. No. 1.2e-106;

Matches 220; Conservative 65; Mismatches 75; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGGVCGVITSGCPRIPLTEADLOVELDRRRQCSRITSTRKET 112  
 Db 2 GNIFGHLFRISTFGESHGGVGVWIDGCPPOLEISAEIQVELDRRRPQSKITTPRKEA 61

QY 113 DTCKILSGTHEGVTGTPIVVPNTDQIGSDHREIANVYRPSHADATYDFKVGVRVQ 172  
 Db 62 DTCEILSGVFEKTLGTPITILVRNQDTRPDQYDEMAQKVRPSHADATYDAKYGRNWQ 121

QY 173 GRRSGRKTGVRVAAGALPKKILKLCGLLEILSFVSKVHVVLPEADYGVSTLEQIES 232  
 Db 122 GGRSSARETIGRVAAGAIAKKILRQVANVEIYVKRIKDL---EGVDPNTVTLDOVES 178

QY 233 NIVRCPDPEYAEKMDAIDRVVRGDSVGVITCVARNVPRGLGSPVFDKLESELAKAML 292  
 Db 179 NIVRCPDAECGRMIELIEIQGQDSIGGVVCEVARNVPGKLGEPVFDKLEADIADAKVM 238

QY 293 SIPASNGFEIGSGFAGTDLTGSEHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 352  
 Db 239 SLPASRGFEIGSGFAGTLLTGFEHNDYIDENGEITRTVNRSGGICQGIANGENIILRV 298

QY 353 AFKPTPSIGVKONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQDMAHVAQ 412  
 Db 299 AFKPTATIRKEQKTVTRGEETLLAAKGRHDPVLPRAVPVMEAMVALVLCDDLRLHHQ 358

QY 413 CEM 415

Db 359 CKV 361

## RESULT 14

AAG24019  
 ID AAG24019 standard; protein; 313 AA.

XX AAG24019;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 27536.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
PP	25-FEB-2000; 2000EP-00301439.	PR	09-JUL-1999;	99US-0142920P.
XX		PR	12-JUL-1999;	99US-0142977P.
PR	25-FEB-1999; 99US-0121825P.	PR	13-JUL-1999;	99US-0143542P.
PR	05-MAR-1999; 99US-0123180P.	PR	14-JUL-1999;	99US-0143624P.
PR	09-MAR-1999; 99US-0123548P.	PR	15-JUL-1999;	99US-0144005P.
PR	23-MAR-1999; 99US-0125788P.	PR	16-JUL-1999;	99US-0144085P.
PR	25-MAR-1999; 99US-0126264P.	PR	16-JUL-1999;	99US-0144086P.
PR	29-MAR-1999; 99US-0126789P.	PR	19-JUL-1999;	99US-0144325P.
PR	01-APR-1999; 99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999; 99US-0128234P.	PR	19-JUL-1999;	99US-0144332P.
PR	08-APR-1999; 99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999; 99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999; 99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999; 99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999; 99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999; 99US-0130891P.	PR	20-JUL-1999;	99US-0144684P.
PR	28-APR-1999; 99US-0131444P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999; 99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999; 99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999; 99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999; 99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999; 99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999; 99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999; 99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999; 99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999; 99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999; 99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999; 99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999; 99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999; 99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999; 99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999; 99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999; 99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999; 99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999; 99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999; 99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999; 99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999; 99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999; 99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999; 99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999; 99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999; 99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999; 99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999; 99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999; 99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999; 99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999; 99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999; 99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999; 99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999; 99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999; 99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999; 99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999; 99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999; 99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999; 99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999; 99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999; 99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999; 99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999; 99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999; 99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999; 99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999; 99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999; 99US-0140353P.	PR	31-AUG-1999;	99US-0151338P.
PR	23-JUN-1999; 99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999; 99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999; 99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	30-JUN-1999; 99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999; 99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999; 99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999; 99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999; 99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.

PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157119P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158023P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159323P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity

Best Global Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0

Qy	125	VTTCPTILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGRAVQGGSSGRKTVGR	184
Db	1	MTTGPIHVFPVNTDQIGLDYSENSVAYRPSHADATYDMKYGRAVQGGSSARETIGR	60
Qy	185	VAAQALPKTKILKCGLEILSFVSKVHOVLPEDAVDYGSVTLEQIESNIYRCDDPEYAE	244
Db	61	VAFALAKKILKQFAGTEILAYVSQVHHVLPPEELVDHENITLEQIENNIYRCNPPEYAE	120
Qy	245	KMIDAI DRVVRVGRDGSVGGVITCVARNVPRGLGSPVDFKLESELAKAMLSIPASNGFEIGS	304
Db	121	KMIAAIDAVRTKGSVGGVITCIYRNA PRGLGTPVDFDKLEAE LAKACMSLPATKGFEGS	180
Qy	305	GFAGTDLTGSEHNDEFYMDKAGSVRTFTNRSGGVQGGISNVEIVHFVKVAFKPTPSIGVKQ	364
Db	181	GFAGTFLTGLHNDEFYTDENGRI TRTNRSGGIQGGISNGEINMRVAFKPTSTIGRKQ	240
Qy	365	NTVSRERQNVLELLARGHDPCCVAPRAVPVSESMAALVMDQLMAHVACCEMFALNTALQE	424
Db	241	NTVTRDQVETENIARGHDPCCVPRAPVPSVSWALVLDVQDMAQYAOCHLFFINPELQE	300
Qy	425	PV	426
Db	301	PL	302

RESULT 15  
 AAY44892  
 ID AAY44892 standard; protein; 257 AA.  
 XX  
 AC AAY44892;

XX		18-MAY-2000 (first entry)	
DT		Wheat chorismate synthase encoded by clone wreln.pk0094.e6.	
XX			
DE		Wheat cDNA clone wreln.pk0094.e6; chorismate synthase;	
XX		immunological screening; herbicide resistance; antibody; gene mapping;	
KW		wheat.	
XX			
OS		Triticum aestivum.	
XX			
PV		WO200005353-A2.	
PN			
XX		03-FEB-2000.	
PD			
XX			
FF		20-JUL-1999; 99WO-US016353.	
XX			
PR		21-JUL-1998; 98US-0093611P.	
PR		(DUPO ) DU PONT DE NEMOURS & CO E I.	
PA			
XX		Cahoon RE, Falco SC, Pember SO;	
PI			
XX		WPI; 2000-182687/16.	
DR		N-PSDB; AAZ50437.	
DR			
XX		New chorismate synthase polypeptides used to alter the level of the enzyme and thus the level of aromatic to non aromatic amino acids in transformed plants.	
PT			
PT		Claim 6; Page 31-32; 39pp; English.	
XX			
PS		The present sequence is wheat chorismate synthase encoded by cDNA clone designated wreln.pk0094.e6. Manipulating either the amount or activity of this enzyme would affect the ratio of aromatic to non-aromatic amino acids in plants. Chimeric gene comprising the cDNA operably linked to regulatory sequences is used to transform host cells to alter the level of expression of chorismate synthase. The gene and its products may be used for immunological screening of cDNA expression libraries and to create transgenic plants which may also be herbicide resistant. Synthetic peptides derived from the gene are to raise antibodies, and used in screening assays to identify inhibitors which may be useful as herbicides	
CC			
XX		Sequence 257 AA;	
SQ			
Query Match 48.5%; Score 1066; DB 3; Length 257;			
Best Local Similarity 82.7%; Pred. No. 1.8e-95;			
Matches 201; Conservative 27; Mismatches 15; Indels 0; Gaps 0;			
Qy	186	AAGALPKKILKLCGLLEILSVSKVHQVLPEDAVDYGSVTLEOIESNIVRCPPDYAEK	245
Dd	3	ARGAVKKILLKLCGVEILAFVSVKHQVLPEDAVDYETLTLDQIESNICRCPDPYAAQK	62
Qy	246	MIDAIDRVVRGDSSVGGVITCVARNVPRGLGSPVFDKLESELAKAMLSIPASNGFEIGSG	305
Dd	63	MIDAIDKRVNGNSICGGVTCIARNVPRGLGSPVFDKLEALLAKAMLSLPASKGFEIGSG	122
Qy	306	FAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVETVHKFAFKPTTPSIGVKQN	365
Dd	123	FAGTDLTGSEHNDEFYMDAEGNVRTNRNSRGVGGISNGETIYFKVAFKPTATIGKKQN	182
Qy	366	TVSREQRONVELLARGHRDPCVAPRAVPVVSMAALVLMDOLMHVAOCFPMALTALQEP	425
Dd	183	TVTREDIEDIELTFGRHDPCVPRAPVPMVETMAALVLMDOQLMHAVQCEMPFLNALQEP	242
Qy	426	VGS 428	
Dd	243	IGS 245	

Search completed: August 25, 2005, 19:30:06  
Job time : 80 secs

**THIS PAGE BLANK (USPTO)**



## RESULT 2

S40410  
Chorismate synthase (BC 4.2.3.5) 1 precursor - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 19-May-1994 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S40410; S38470  
R;Goerlach, J.; Schmid, J.; Amrhein, N.  
Plant Mol. Biol. 23, 707-716, 1993  
A;Title: Differential expression of tomato (Lycopersicon esculentum L.) genes encoding a  
A;Reference number: S40409; MUID:94072719; PMID:8251624  
A;Accession: S40410  
A;Molecule type: mRNA  
A;Residues: 1-440 <GO>  
A;Cross-references: UNIPROT:Q42884; EMBL:Z21796; NID:G410481; PIDN:CAA79859.1; PID:G4104

C;Genetics:  
A;Genome: nuclear  
C;Superfamily: chorismate synthase  
C;Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; chloroplast; phosphon  
F;1-54/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F;55-440/Product: chorismate synthase 1 #status predicted <MAT>

Query Match 69.4%; Score 1527; DB 2; Length 440;  
Best Local Similarity 67.4%; Pred. No. 3.4e-107;  
Matches 288; Conservative 64; Mismatches 73; Indels 2; Gaps 1;

Qy 4 VPKQFQVAHSRAR--LAPRAIGALLEFAPAPASSLRFAVHRCRTARLEVKASGNTFGNYFQ 61

Db 5 VPTKQFVGASSSDIGSSRLVSLQLPSKFSSNFHLPSPQSLKRLKLEIQAAAGSTFGNYFR 64

Qy 62 VATTGESGGGVCVIGSCPPRIPLTEADLOVELDRRPPGOSRITSTRKETDTCKILSGT 121

Db 65 VTTGESGGGVCIIIDCCPRLPPLPSESDMQVELDRRPPGOSRITSTRKETDTCKILSGT 124

Qy 122 HEGVTTPFPIIVPNTDQIGSDHREIANVYRPSHADATYFKYGVRAVQGGSSGRKT 181

Db 125 ADGLTTGSPKVEPNTDQNDYSEMSLAYRPSHADATYFKYGVRAVQGGSSGRKT 184

Qy 182 VGRVAAGALPKILKLCGLSILFSVKVHQVLPEDAVYGVSTVLEQIESNIVRCDDPE 241

Db 185 IGRVAAGAVAKILKLYSGAEVLAVYSQVHQVLPEDLDHQNTLEQIESNIVRCDDPE 244

Qy 242 YAEKMDAIDRVVRGDSVGVITCVARNVRGLSGSPVFDKLESELAKAMLSIASNGFE 301

Db 245 YAEKMDAIDRVVRGDSVGVVITCVARNVRGLSGTPVFDKLESELAKAMLSIASNGFE 304

Qy 302 IGSFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQGGVGSINVEIVHFKVAFKPTPSIG 361

Db 305 FGSFAGTDMTGSEHNDEFYMDKAGSVRTRNRSQGGVGSINVEIVHFKVAFKPTPSIG 364

Qy 362 VKQNTVSRERQNVELLARHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTA 421

Db 365 RKQQTVRTDKHETELIARGHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTA 424

Qy 422 LQEPVGS 428

Db 425 LQEPVGS 431

RESULT 3

A41197  
Chorismate synthase (BC 4.2.3.5) precursor [validated] - pink corydalis  
C;Species: Corydalis sempervirens (pink corydalis)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: A41197; S13267; S15715  
R;Schaller, A.; Schmid, J.; Leibinger, U.; Amrhein, N.  
J. Biol. Chem. 266, 21434-21438, 1991  
A;Title: Molecular cloning and analysis of a cDNA coding for chorismate synthase from th  
A;Reference number: A41197; MUID:92402037; PMID:1718979  
A;Accession: A41197  
A;Status: preliminary  
A;Molecule type: mRNA

A;Residues: 1-447 <SCH>

A;Cross-references: UNIPROT:P27793; EMBL:X60544; NID:gl8255; PIDN:CAA43034.1; PID:gl8255  
R;Schaller, A.; Windhofer, V.; Amrhein, N.  
Arch. Biochem. Biophys. 282, 437-442, 1990

A;Title: Purification of chorismate synthase from a cell culture of the higher plant Cory  
A;Reference number: S13267; MUID:91053166; PMID:2146922

A;Accession: S13267

A;Molecule type: protein

A;Residues: "X", 307-320 <SCW>

C;Genetics:

A;Genome: nuclear

C;Complex: homodimer

C;Function:

A;Description: EC 4.6.1.4 [validated, MUID:91053166]; catalyzes the formation of chorism  
C;Superfamily: chorismate synthase  
C;Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; chloroplast; phosphon  
F;1-57/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F;58-447/Product: chorismate synthase #status predicted <MAT>

Query Match 67.6%; Score 1486.5; DB 1; Length 447;  
Best Local Similarity 70.4%; Pred. No. 3.8e-104;  
Matches 280; Conservative 60; Mismatches 55; Indels 3; Gaps 2;

Qy 31 ASSSLRPAVHRCRTA--RLEVKAAGNTFGNYFQVATYGESHGCGVGVISGCCPPRIPLTE 88

Db 37 SNQSVQISVPR-QTAPLKLVVQASGSGFGKVFQVTTYGESHGCGVGVISGCCPPRIPISE 95

Qy 89 ADIQVELDRRRPGOSRITSTRKETDTCKILSGTHEGVTTPILVIVPNTDQIGSDHREI 148

Db 96 ADIQVELDRRRPGOSRITSTRKETDTCKILSGVADGFTTGTSPHISVPNTDQIGSDHREI 155

Qy 149 ANVYRPSHADATYFKYGVRAVQGGSSGRKTGRVAAGALPKILKLCGLSILFSVS 208

Db 156 AKYRPSHADATYFKYGVRAVQGGSSGRKTGRVAAGALPKILKLCGLSILFSVS 215

Qy 209 KVHQVLPEDAVYGVSTVLEQIESNIVRCDDPEYAEKMDAIDRVVRGDSVGVITCV 268

Db 216 QAHKVLPEGLVDHETLSLEQIESNIVRCDDPEYAEKMDAIDRVVRGDSVGVITCV 275

Qy 269 RNVPRLGSGVPFKLSESELAKAMLSIASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSV 328

Db 276 RNVPRLGSGVPFKLSESELAKAMLSIASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSV 335

Qy 329 RTNRSQGGVGSINVEIVHFKVAFKPTPSIGVKQNTVSRERQNVELLARHDPVAP 388

Db 336 RTNRSQGGVGSINVEIVHFKVAFKPTPSIGVKQNTVSRERQNVELLARHDPVAP 395

Qy 389 RAVPVVSMAALVMDQLMAHVAQCEMPALNTALQEPV 426

Db 396 RAVPVVSMAALVMDQLMAHVAQCEMPALNTALQEPV 433

RESULT 4

C96526

probable chorismate synthase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: C96526

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C96526

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-435 <STO>



A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <KAN>  
A:Cross-references: EMBL:D90906; GB:AB001339; NID:g1652492; PIDN:BAAL7415.1; PID:g1652492  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: aroC  
C:Superfamily: chorismate synthase  
C:Keywords: carbon-oxygen lyase; phosphorus-oxygen lyase

Query Match 54.2%; Score 1191.5; DB 2; Length 362;  
Best Local Similarity 62.2%; Pred. No. 4.5e-82;  
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGGSHGGVGCIVISGCPPIRLPLTEADLQVELDRRPGOSRITSTRKET 112  
DB 2 GNTFGSLFRITTFGSHGGVGVIIIDGCPRLISPEEIQVLDLDRRPGQSKITTPRKEA 61

QY 113 DTCKILSGTHEGVTGTPTLIVIVNTDQIGSDHREITANVYRPSHADATYDFKYGVRAVQG 172  
DB 62 DQCEILSGVFEGKTLGTPTAILVRNKDARSQDYNEMAVKYRPSHADATYEAKYGI RNWQ 121

QY 173 GRRSGRKTGVRAAGALPKKILKCGLEILLSFVSKVQHVLPEDAVDYGSVTLQIES 232  
DB 132 GRRSARETIGRAAGAIKALQAPNGVEIVAVKSIQDI---EATVDSNTVTLEQVES 178

QY 233 NIVRCPDPEYAEKMDAIDRVRVRGSGVGVITCVARNVPRGLGSPVFDKLESELAKAML 292  
DB 179 NIVRCPDECAEKMERIDQVLRQKDSIGGVVECAIRNAPKGLGEPVFDKLEADLAKAMM 238

QY 293 SIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNTRSGGVGGISNVBEIVHPKV 352  
DB 239 SLPATKGFEGSGFAGTLLTGSHNDDEYLLDEAGEWKTRNTRSGGVGGISNGSPPIIMRI 298

QY 353 AFKPTPSGVKQNTVSRERONVELLARGHPDCAVRAVPPVSVESMAALVLMQDLMAHVAQ 412  
DB 299 AFKPTATIGQEOKTVSNIGSEETTLAAKGRHDPVLPRAVPMVEAMALVLCDDLRFQAAQ 358

QY 413 CE 414  
DB 359 CK 360

RESULT 6  
AC1906  
chorismate synthase [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AC1906  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 203-213, 2001  
A:Title: Complete Genomic  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AC1906  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <KUR>  
A:Cross-references: UNIPROT:Q8YTP9; GB:BA000019; PIDN:BAE72754.1; PID:g17130142; GSPDB:G  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all0797  
C:Superfamily: chorismate synthase

Query Match 53.9%; Score 1186.5; DB 2; Length 362;  
Best Local Similarity 60.6%; Pred. No. 1.1e-81;  
Matches 220; Conservative 66; Mismatches 74; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGGSHGGVGCIVISGCPPIRLPLTEADLQVELDRRPGOSRITSTRKET 112  
DB 2 GNTFGHLFRITTFGSHGGVGVIIIDGCPPLISPEEIQLDLDRRPGQSKITTPRKEA 61



QY 406 LMAHVAQ 412  
| : :  
Db 362 LLIQXAR 368

## RESULT 9

T46725  
Chorismate synthase (EC 4.2.3.5) / flavin reductase, NADPH-dependent [validated] - Neuro  
N:Contains: chorismate synthase (EC 4.6.1.4); flavin reductase, NADPH-dependent  
C:Species: Neurospora crassa  
C>Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: T46725  
R:Henstrand, J.M.; Amrhein, N.; Schmid, J.  
J. Biol. Chem. 270, 20447-20452, 1995  
A:Title: Cloning and Characterization of a Heterologously Expressed Bifunctional Chorism  
A:Reference number: Z08450; MUID:95386486; PMID:7657620  
A:Accession: T46725  
A:Molecule type: mRNA  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Residues: 1-434 <HEN>  
A:Cross-references: UNIPROT:Q12640; EMBL:U25818; NID:g976374; PIDN:AAC49056.1; PID:g9763

## C:Function:

A:Description: EC 4.6.1.4 [validated, MUID:95386486]  
A:Note: N. crassa chorismate synthase (CS) is a bifunctional enzyme, possessing also an  
A:Note: CS and FT activity are encoded by a single domain  
C:Function: <PRE>  
A:Description: Flavin reductase [validated, MUID:95386486]  
A:Note: N. crassa chorismate synthase (CS) is a bifunctional enzyme, possessing also an  
A:Note: CS and FT activity are encoded by a single domain  
C:Superfamily: chorismate synthase  
C:Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; phosphorus-oxygen ly

Query Match 42.3%; Score 930; DB 2; Length 432;

Best Local Similarity 48.9%; Pred. No. 2.7e-62;

Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;

QY 54 NTFGNYFOVATYGESHGGVGVISGCPRIPLTEADLQVELDRRPPGOSRITSTRKETD 113

Db 2 STFGHYFRVTYTGESHCKSVGCVVDGVPVPGMELTDDIQPMTRRRPGOSAITTPRDEKD 61

QY 114 TKILSGTHEGTGTPILVIVPNTDQIGSDH-REIANVY-RPSHADATYDFKYGVRVQ 171

Db 62 RVLIQSGTEFGVTGLTPIGMLVMNEDQPKDYGNKTMIDYPRPSHADWTYLEKYGKASS 121

QY 172 GGRSSGRKTGVRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED 218

Db 122 GGRSSARETIGRVAAGAIAEKYLKPRYGVVEIVAFVSSVSGSEHLFPPTAEHPSPSTNPEF 181

QY 219 AVDYGSVTLEQIESNI-VRCPOPEYAEKMDAIDRVRVSGSVGVITCVARNVPRGLGS 277

Db 182 LKLVNSITRETVDSPFVRCPDAEANKMEDLITKFRDNHDSIGGVITCVIRNVPGLGE 241

QY 278 PVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327

Db 242 PAFDKLEAMLAHMLSIPTATKFEVSGGCGCEVPGSIHNDPFVSAENTEIPPSVAASGA 301

QY 328 VR-----TRTNRSGVQGGISNVEIVHFVKAFKPTPSIGVKONTVSRERQNVLL- 377

Db 302 ARNGIPRPKLTKTNFSGGIQGGISNGAPIYFRVGFKPAATIGOEQTATYDGTSGVLA 361

QY 378 ARGHRDPCVAPRAVPVVESSMAALVMDQLMAHVAOCFALNTALQEPVGS 428

Db 362 AKGRHDPSPVPRAPVIVEAMALVMDAIVLAHEARVTAKSLPPLKQTINS 412

## RESULT 10

T51020  
Chorismate synthase/flavin reductase, NADPH-dependent [imported] - Neurospora crassa  
N:Alternate names: protein B7F21.10  
C:Species: Neurospora crassa  
C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
C:Accession: T51020

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, July 2000  
A:Reference number: Z25286

A:Accession: T51020

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-432 <SCH>

A:Cross-references: EMBL:AL389901; GSPDB:GN00116; NCSP:B7F21.10

A:Experimental source: BAC clone B7F21; strain OR74A

C:Genetics:

A:Gene: NCSP:B7F21.10

A:Map position: 6

A:Introns: 12/2; 92/1

C:Superfamily: chorismate synthase

Query Match 42.0%; Score 924; DB 2; Length 432;

Best Local Similarity 48.9%; Pred. No. 7.6e-62;

Matches 201; Conservative 63; Mismatches 111; Indels 36; Gaps 7;

QY 54 NTFGNYFOVATYGESHGGVGVISGCPRIPLTEADLQVELDRRPPGOSRITSTRKETD 113

Db 2 STFGHYFRVTYTGESHCKSVGCVVDGVPVPGMELTDDIQPMTRRRPGOSAITTPRDEKD 61

QY 114 TKILSGTHEGTGTPILVIVPNTDQIGSDH-REIANVY-RPSHADATYDFKYGVRVQ 171

Db 62 RVLIQSGTEFGVTGLTPIGMLVMNEDQPKDYGNKTMIDYPRPSHADWTYLEKYGKASS 121

QY 172 GGRSSGRKTGVRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED 218

Db 122 GGRSSARETIGRVAAGAIAEKYLKAYGVVEIVAFVSSVSGSEHLFPPTAEHPSPSTNPEF 181

QY 219 AVDYGSVTLEQIESNI-VRCPPPEYAEKMDAIDRVRVSGSVGVITCVARNVPRGLGS 277

Db 182 LKLVNSITRETVDSPFVRCPDAEANKMEDLITKFRDNHDSIGGVITCVIRNVPGLGE 241

QY 278 PVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327

Db 242 PAFDKLEAMLAHMLSIPTATKFEVSGGCGCEVPGSIHNDPFVSAENTEIPPSVAASGA 301

QY 328 VR-----TRTNRSGVQGGISNVEIVHFVKAFKPTPSIGVKONTVSRERQNVLL- 377

Db 302 ARNGIPRPKLTKTNFSGGIQGGISNGAPIYFRVGFKPAATIGOEQTATYDGTSGVLA 361

QY 378 ARGHRDPCVAPRAVPVVESSMAALVMDQLMAHVAOCFALNTALQEPVGS 428

Db 362 AKGRHDPSPVPRAPVIVEAMALVMDAIVLAHQARHTAKSLPPLKQTINS 412

## RESULT 11

AG0335

Chorismate synthase (EC 4.2.3.5) [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C:Accession: AG0335

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

ii, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0335

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <KUR>

A:Cross-references: UNIPROT:Q8ZD41; GB:AL590842; PIDN:CAC92990.1; PID:g15980729; GSPDB:

C:Genetics:

A:Gene: aroC

C:Superfamily: chorismate synthase

C:Keywords: carbon-oxygen lyase; phosphorus-oxygen lyase

Query Match 39.2%; Score 862; DB 2; Length 361;

Best Local Similarity 48.3%; Pred. No. 2.8e-57;

Matches 175; Conservative 65; Mismatches 110; Indels 12; Gaps 3;



```
Db 122 GGRSSARETAMRVAAGAAIAKYLAEKFGIEIRGCLTQMGGDI--PLEIKDW-----SQVE 174
QY 232 SNIVRCPDPEYAEKMDAIDR----VRVRGDSVGGVITTCVARNVPRGLGSPVFDKLESEL 287
Db 175 QNPFPCPDPD----KIDALDELMRALKKEGDSIGAKVTIVVAGVPAGLGEVPFDRDLADI 230
QY 288 AKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQGVGGISNVBI 347
Db 231 AHALMSINAVKVGVEIGDGFVVALRGSQNRDEITKD-----GFQSNHAGGILGGISSGQQ 285
QY 348 VHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPVAPRAVPVVSMAALVMDQLM 407
Db 286 IIAHMAKPTSSITVPGRTINRFGEEVEMITKGRHDPVCGIRAVPIAEAMLAIVLMDHLL 345
QY 408 AHVAQ 412
Db 346 RQRAQ 350

RESULT 14
F85874
Chorismate synthase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85874
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85874
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <STO>
A:Cross-references: UNIPROT:Q8XCQ4; GB:AE005174; NID:ig12516691; PIDN:AA657458.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: aroC
C:Superfamily: chorismate synthase

Query Match 38.9%; Score 856; DB 2; Length 361;
Best Local Similarity 48.5%; Pred. No. 7.8e-57;
Matches 177; Conservative 62; Mismatches 106; Indels 20; Gaps 5;

QY 52 SGNRTGNFQVATYGESHGCGVGVCGPPIPLTEADLQVELDRRRPGQSRTSTKE 111
Db 2 AGNTIGQLFRVITTFGESHGLAGCIVDGVPPGIPTEADLQHDLDLDRRRPGTSRYTTORRE 61
QY 112 TDCKILSGTHGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDKYGVRAVQ 171
Db 62 PDQVILSGVFGVITGTSIGLLIENTDQRSQDYSNAIKVDVFRPGHADYTYEQKYLDRYR 121
QY 172 GGRSSGRTVGRVAGALPKILKLCGLLEILSFVSKVHQVLPEDAVDYGVSVTLEQIE 231
Db 122 GGRSSARETAMRVAAGAAIAKYLAEKFGIEIRGCLTQMGGDI--PLEIKDW-----SQVE 174
QY 232 SNIVRCPDPEYAEKMDAIDR----VRVRGDSVGGVITTCVARNVPRGLGSPVFDKLESEL 287
Db 175 QNPFPCPDPD----KIDALDELMRALKKEGDSIGAKVTIVVAGVPAGLGEVPFDRDLADI 230
QY 288 AKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQGVGGISNVBI 347
Db 231 AHALMSINAVKVGVEIGDGFVVALRGSQNRDEITKD-----GFQSNHAGGILGGISSGQQ 285
QY 348 VHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPVAPRAVPVVSMAALVMDQLM 407
Db 286 IIAHMAKPTSSITVPGRTINRFGEEVEMITKGRHDPVCGIRAVPIAEAMLAIVLMDHLL 345
QY 408 AHVAQ 412
Db 346 RQRAQ 350
```

```
RESULT 15
A55510
Chorismate synthase (EC 4.2.3.5) - Vibrio anguillarum
C:Species: Vibrio anguillarum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jun-2002
C:Accession: A55510
R:Chen, Q.; Actis, L.A.; Tolmasky, M.E.; Crosa, J.H.
J. Bacteriol. 176, 4226-4234, 1994
A:Title: Chromosome-mediated 2,3-dihydroxybenzoic acid is a precursor in the biosynthesis
A:Reference number: A55510; MUID:94292450; PMID:8021209
A:Accession: A55510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <CHE>
A:Cross-references: GB:L29562
C:Genetics:
A:Gene: aroC
C:Superfamily: chorismate synthase
C:Keywords: carbon-oxygen lyase; phosphorus-oxygen lyase

Query Match 38.6%; Score 850; DB 1; Length 372;
Best Local Similarity 46.0%; Pred. No. 2.3e-56;
Matches 172; Conservative 70; Mismatches 116; Indels 14; Gaps 4;

QY 47 LEVKASGNTFNGYFQVATYGESHGCGVGVCGPPIPLTEADLQVELDRRRPGQSRT 106
Db 8 IEEVWAGNSIQHFRVMTFGESHGIALGCIVDGCPGLEITEADLQIDLDRRRPGTSRYT 67
QY 107 STRKETDTCKILSGTHGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKY 166
Db 68 TORREADEVKILSGVFEKTTGTSGILLIENTDQRTDYSIDKDKFRPGHADYTYHQYK 127
QY 167 VRVQGGGSGRKTGVRVAGALPKILKLCGLLEILSFVSKVHQVLPEDAVDYGSVT 226
Db 128 IRDYRGGGSRARETAMRVAAGAAIAKYLKQEFVETRAYLSQMGDVCI--DKVDW---- 181
QY 227 LEQIESNIVRCPDPEYAEKMDAIDRVRVRGDSVGGVITTCVARNVPRGLGSPVFDKLESE 286
Db 182 -NEIENNAFFCPDADKVAAPDQLIRDLKKEGDSIGAKIQVATNLVGLGEPVFDRLDAD 240
QY 287 LAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQGVGGISNVBI 346
Db 241 IAHALMSINAVKVGVEIGDGFVQVQKGSQHRDLPTN-----GFRSNHAGGILGGISTGQ 295
QY 347 IVHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPVAPRAVPVVSMAALVMDQL 406
Db 296 DIVASIALKPTSSITVPGDTITRTGPTQLITKGRHDPVCGIRAVPIAEAMLAIVLMDHL 355
QY 407 MAHVAQCCEMFALNT 420
Db 356 LRHRGQ--NFAVQT 367

Search completed: August 25, 2005, 19:31:39
Job time : 20 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:22:15 ; Search time 67 Seconds  
(without alignments)  
3278.835 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1541	70.0	431	1	ARC2_LYCES	Q42885 lycopersico
2	1527	69.4	440	1	ARC1_LYCES	Q42884 lycopersico
3	1486.5	67.6	447	1	AROC_CORSE	P27793 corydalis s
4	1454	66.1	436	2	Q93Z13	Q93zi3 arabidopsis
5	1438.5	65.4	435	1	AROC_ARATH	P57220 arabidopsis
6	1191.5	54.2	362	1	AROC_SYNY3	P23353 synchocyst
7	1186.5	53.9	362	1	AROC_ANASP	Q81m1 synchococcc
8	1176.5	53.5	363	2	Q8DLM1	Q8dlm1 synchococcc
9	1047	47.6	364	2	Q7V364	Q7v364 prochloroco
10	1038	47.2	363	2	Q7U9F0	Q7u9f0 synchococcc
11	1023.5	46.5	362	1	AROC_PROWA	P46894 prochloroco
12	1007.5	45.8	362	2	Q7V4Y9	Q7v4y9 prochloroco
13	951	43.2	395	1	AROC_SCHPO	Q74413 schizosacch
14	950.5	43.2	382	2	Q6AIF3	Q6aip3 deulifotale
15	944	42.9	376	2	Q6CNY1	Q6cny1 kluyveromyc
16	944	42.9	376	2	Q8FV38	Q8fv38 candida gla
17	943	42.9	376	2	Q7S5B9	Q7s5b9 ashbya gos
18	938	42.6	376	1	AROC_YEAST	P28777 debaryomyce
19	938	42.6	377	2	Q6BQN1	Q6bqn1 debaryomyce
20	928.5	42.2	380	2	Q72W01	Q72w01 leptospira
21	928.5	42.2	380	2	Q8F9N4	Q8f9n4 leptospira
22	924.5	42.0	375	2	Q6C8Q1	Q6c8q1 yarrowia li
23	924	42.0	432	1	AROC_NEUCR	Q12640 neurospora
24	889.5	40.4	358	2	Q64PF7	Q64pp7 bacteroides
25	887	40.3	219	2	Q75W14	Q75w14 oryza sativ
26	874.5	39.8	358	2	Q8A602	Q8a602 bacteroides
27	867	39.4	367	2	Q649I0	Q649i0 uncultured
28	863	39.2	361	2	Q6LNN4	Q6lnn4 photobacter
29	862	39.2	360	1	AROC_YERPE	Q8xd41 yersinia pe
30	862	39.2	361	2	Q668V5	Q668v5 yersinia ps
31	860	39.1	360	1	AROC_ECOLI	P12008 escherichia

32 858 39.0 361 1 AROC\_VIBPA Q87mm9 vibrio para  
33 856 38.9 360 1 AROC\_ECO57 P63610 escherichia  
34 856 38.9 360 1 AROC\_ECOL6 P63609 escherichia  
35 853 38.8 361 1 AROC\_VIBAN P39198 vibrio angu  
36 851 38.7 361 2 Q6D2M6 Q6d2m6 erwinia car  
37 849 38.6 361 2 Q7UC49 Q7uc49 shigella fl  
38 849 38.6 362 2 Q83Q05 Q83qq5 shigella fl  
39 847.5 38.5 377 2 Q7MIT1 Q7mit1 vibrio vuln  
40 846.5 38.5 357 1 AROC\_PASMU P57840 pasteurella  
41 845 38.4 361 1 AROC\_VIBVU Q8db42 vibrio vuln  
42 845 38.4 365 1 AROC\_METWA Q8pw84 methanosarc  
43 844 38.4 353 1 AROC\_BUCAP Q9zhe9 buchnera ap  
44 843 38.3 360 1 AROC\_SALTY P58729 salmoneila  
45 843 38.3 365 1 AROC\_METAC Q8tt87 methanosarc

#### ALIGNMENTS

RESULT 1  
ARC2\_LYCES STANDARD; PRT; 431 AA.  
AC Q42885;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Chorismate synthase 2, chloroplast precursor (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase 2).  
DE enolpyruvylshikimate-3-phosphate phospholyase 2).  
GN Name=CS2;  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. UC82B;  
RX MEDLINE=94072719; PubMed=8251624;  
RA Georlach J., Schmid J., Anrheim N.;  
RT "Differential expression of tomato (Lycopersicon esculentum L.) genes encoding shikimate pathway isoenzymes. II. Chorismate synthase.";  
RL Plant Mol. Biol. 23:707-716(1993).  
CC -!- FUNCTION: Catalyzes the last common step of the biosynthesis of aromatic amino acids, produced via the shikimic acid pathway.  
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate = chorismate + phosphate.  
CC -!- COFACTOR: Reduced flavin.  
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; seventh step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Predominantly expressed in flowers and roots and, to a lesser extent, in stems, leaves, and cotyledons.  
CC -!- TISSUE SPECIFICITY: Belongs to the chorismate synthase family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; Z21791; CAA79854.1; -;  
PIR; S40409; S40409.  
DR InterPro; IPR000453; Chorismate synth.  
DR Pfam; PF01264; Chorismate synt; 1.  
DR ProDom; PD002941; Chorismate synth; 1.  
DR TIGRFAMs; TIGR00033; aroc\_1\_1.  
DR PROSITE; PS00787; CHORISMATE SYNTHASE\_1; 1.  
DR PROSITE; PS00788; CHORISMATE SYNTHASE\_2; 1.  
DR PROSITE; PS00789; CHORISMATE SYNTHASE\_3; 1.  
KW Aromatic amino acid biosynthesis; Chloroplast; Lyase; Transit peptide.





Qy 329 RRTNRSGVQGGISNVEIVHKFAFKTPTSGVQNTVSRQRNVELLARGHRDPCVP 389  
 Db 336 RRTNRSGIQGGISNGEIIINRIAFKPTSTIGKKQNTVTREREETELIARGHRDPCVP 395  
 Qy 389 RAVPVVESNALVLMQDLMAHVAQCEPALNTALQEPV 426  
 Db 396 RAVPVVEAWVALVLDQLMLQHAQGNLFSINPALQEPV 433

RESULT 4  
 Q93Z13

ID	Q93Z13	PRELIMINARY;	PRT;	436 AA.
AC	Q93Z13;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	At1648950/724P22.3.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eucoids II; Brassicales; Brassicaceae; Arabidopsi.			
NCBI	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Chen R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,			
RA	Bowser L., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,			
RA	Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,			
RA	Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,			
RA	Nguyen M., Onodera C.S., Palm C.J., Phan P.K., Quach H.L., Sakurai T.,			
RA	Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,			
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,			
RA	Ecker J.R.;			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =			
CC	chorismate + phosphate.			
CC	-1- COFACTOR: Reduced flavin (By similarity).			
CC	-1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;			
CC	seventh step.			
CC	-1- SIMILARITY: Belongs to the chorismate synthase family.			
DR	EMBL; AY057519; AAL09759.1; --			
DR	HSSP; P28777; 1R52.			
DR	GO; GO:0004107; F:chorismate synthase activity; IEA.			
DR	GO; GO:0016829; F:lyase activity; IEA.			
DR	GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.			
DR	InterPro; IPR000453; Chorismate synth.			
DR	Protein; IPRO11013; Gal mut like.			
DR	Pfam; PF01264; Chorismate_synth; 1.			
DR	ProDom; PD002941; Chorismate_synth; 1.			
DR	TIGRFAMs; TIGR00033; aroc; 1.			
DR	PROSITE; PS00787; CHORISMATE SYNTHASE_1; 1.			
DR	PROSITE; PS00788; CHORISMATE SYNTHASE_2; 1.			
DR	PROSITE; PS00789; CHORISMATE SYNTHASE_3; 1.			
KW	Amino-acid biosynthesis; Aromatic amino acid biosynthesis; Lyase.			
SEQ	SEQUENCE 436 AA; 47317 MW; 6753188EA056680F CRC64;			

Query Match 66.1%; Score 1454; DB 2; Length 436;  
 Best Local Similarity 67.7%; Pred. No. 9.4e-99;  
 Matches 268; Conservative 65; Mismatches 63; Indels 0; Gaps 0

Qy 31 ASSSLRFAVHCRTRARLEVKASNGTNGYFQVATYGESGGVGVCGVSGCPRIPLTEAD 90  
 Db 30 SSPAVQISLRQTQKRFQIQATGSSYGTFRFVSTFGESGGVGCIDGCPRIPLTESD 89  
 Qy 91 LQVLEDRRPPQSRITSTRKETDTCKLSCTGTHSGVTTGTPILVIVPNTDQIGSDHREIAN 150  
 Db 90 LQVLEDRRPPQSRITSTRKETDTCKLSCTGTHSGVTTGTPILVIVPNTDQIGSDHREIAN 149  
 Qy 151 VYRPSHADATYDFKYGVRVQGGGRSSGRKTVGRVAAGALPKKILKLCGLLETSFVSKV 210  
 Db 150 AYRPSHADATYDMKYGVRSVQGGGRSSARETIGRVAFGALAKKILKLCGLLETSFVSKV 209  
 Qy 211 HQVLPDAVDYGSVLTLEQIESNIVRCPDPEYAEKIMDAIDRVVRVGRDVSVGVTTCVARN 270

```
Db 210 HHVLPPELVHDLNLTLEQIENNVRCNPEYAEKMIADAVRTKGSVGGVTCIVRN 269
Qy 271 VPRGLGSPVFDKLESELAAMLSIPASNGFELGSGFAGTDLTGSEHNDEFYMDKAGSVRT 330
Db 270 APRGLGTPVFDKLEAEALAKACMSLPATKGFPGSGFAGTFLTGLEHNDEFYTDENGRI 329
Qy 331 RTNRSGGVQGGISNNVEIVHFVAFKPTSPISGVKQNTVSRQNVVELLARGHDPVAPRA 390
Db 330 RTNRSGGVQGGISNNVEIVHFVAFKPTSPISGVKQNTVSRQNVVELLARGHDPVAPRA 389
Qy 391 VPVVESMAALVMDQLMAHVAQCEMFALNTALQEPV 426
Db 390 VPVVEAMVALVLDQLMAQYAQCHLFPINPELQEP 425

RESULT 5
AROC ARATH STANDARD; PRT; 435 AA.
AC P57720;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Choriarnate synthase, chloroplast precursor (EC 4.2.3.5) (5-
  enolpyruvylshikimate-3-phosphate phospholase).
GN OrderedLocusNames=At1g48850; ORFNames=T24P22.3, F27K7.1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federpiet N.A., Kaul S.,
  White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
  Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
  Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
  Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
  Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
  Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
  Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
  Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
  Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
  Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
  Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
  Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
  Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
  Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
  Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
  "Sequence and analysis of chromosome 1 of the plant Arabidopsis
  thaliana.";
RL Nature 408:816-820(2000).
RC NATURE! Catalyzes the last common step of the biosynthesis of
  aromatic amino acids, produced via the shikimic acid pathway (By
  similarity).
CC -! CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
  chorismate + phosphate.
CC -! COFACTOR: Reduced flavin (By similarity).
CC -! PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
  seventh step.
CC -! SUBUNIT: Homotrimer (By similarity).
CC -! SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -! SIMILARITY: Belongs to the chorismate synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
DR EMBL; AC084242; AAC50662.1; -.
DR EMBL; AC084414; AAC29740.1; -.
DR PIR; C96526; C96526.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRPFAM; TIGR00033; aroC_1.
DR PROSITE; PS00787; CHORISMATE SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE_3; 1.
KW Aromatic amino acid biosynthesis; Chloroplast; Lyase; Transit peptide.
FT TRANSIT 1 49 Chloroplast [Potential].
FT CHAIN 50 435 Chorismate synthase.
SQ SEQUENCE 435 AA; 47161 MW; 822B264B6C5FFC94 CRC64;

Query Match 65.4%; Score 1438.5; DB 1; Length 435;
Best Local Similarity 67.4%; Pred. No. 1.3e-97;
Matches 267; Conservative 65; Mismatches 63; Indels 1; Gaps 1;

Qy 31 ASSSLREAVHRCRTARLEVKASNTFGNYFOVATYGBSHGGVCGVSGCPRIPLTEAD 90
Db 30 SSPAVQISLRTQTRKNPQIQATGSSYGTFRVSTFGSHGGVCGCIIDGCPRIPLTESD 89
Qy 91 LQVELDRRRPQGRITSTRKETDTCKILSGTHGVTGTPILVIVPNTDQIGSDHRIAN 150
Db 90 LQFDLD-RRPQGRITTPKETDTCRISGSGVSGMTGTPIHVFPNTDQGLDYSEMSV 148
Qy 151 VYRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAAGALPKKILKUKGLEILSFVSKV 210
Db 149 AYRPSHADATYDMKYGVRSVQGGSSGRSARETIQVAPGALAKKILKQFAGTEILAYVSQV 208
Qy 211 HQVLPEDAVDYGSVTLEQIESNIVRCPDPEYAEKMIADAVRVRGDSVGGVITCVARN 270
Db 209 HHVLPPELVHDLNLTLEQIENNVRCNPEYAEKMIADAVRTKGSVGGVTCIVRN 268
Qy 271 VPRGLGSPVFDKLESELAAMLSIPASNGFELGSGFAGTDLTGSEHNDEFYMDKAGSVRT 330
Db 269 APRGLGTPVFDKLEAEALAKACMSLPATKGFPGSGFAGTFLTGLEHNDEFYTDENGRI 328
Qy 331 RTNRSGGVQGGISNNVEIVHFVAFKPTSPISGVKQNTVSRQNVVELLARGHDPVAPRA 390
Db 329 RTNRSGGVQGGISNNVEIVHFVAFKPTSPISGVKQNTVSRQNVVELLARGHDPVAPRA 388
Qy 391 VPVVESMAALVMDQLMAHVAQCEMFALNTALQEPV 426
Db 389 VPVVEAMVALVLDQLMAQYAQCHLFPINPELQEP 424

RESULT 6
AROC SYN3 STANDARD; PRT; 362 AA.
AC P23353;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate
  phospholase).
GN Name=aroC; OrderedLocusNames=sl11747;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94086566; PubMed=7505271;
RT "A novel operon organization involving the genes for chorismate
  synthase (aromatic biosynthesis pathway) and ribosomal GTPase center
  proteins (L11, L1, L10, L12: rplKALJ) in cyanobacterium Synechocystis
  PCC 6803.";
RL J. Biol. Chem. 268:27447-27457(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
```

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugiura M., Sasamoto S., Kimura T.,  
RA Hoshouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
RN [3]  
RP SEQUENCE OF: 1-138 FROM N.A.  
RX MEDLINE=91002677; PubMed=2119815; DOI=10.1016/0167-4781(90)90142-O;  
RA Sibold C., Subramanian A.R.;  
RT "Cloning and characterization of the genes for ribosomal proteins L10  
RT and L12 from *Synechocystis* Sp. PCC 6803: comparison of gene clustering  
RT pattern and protein sequence homology between cyanobacteria and  
RT chloroplasts.";  
RL Biochim. Biophys. Acta 1050:61-68(1990).  
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =  
CC chorismate + phosphate.  
CC -!- COFACTOR: Reduced flavin (By similarity).  
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
CC seventh step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SIMILARITY: Belongs to the chorismate synthase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X67516; CAA47855.1; -;  
DR EMBL; D90906; BAA17415.1; -;  
DR EMBL; X53178; CAA37319.1; -;  
DR PIR; A49316; A49316.  
DR HAMAP; MF\_00300; -; 1.  
DR InterPro; IPR000453; Chorismate\_synth.  
DR Pfam; PF01264; Chorismate\_synth; 1.  
DR ProDom; PD002941; Chorismate\_synth; 1.  
DR TIGRFAMs; TIGR00033; aroC; 1.  
DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.  
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.  
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.  
DR Aromatic amino acid biosynthesis; Complete proteome; Lyase.  
SQ SEQUENCE 362 AA; 39287 MW; 9709B0F409168AB8 CRC64;  
Query Match 54.2%; Score 1191.5; DB 1; Length 362;  
Best Local Similarity 62.2%; Pred. No. 1.6e-79;  
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;  
QY 53 GNTFGNYFOVATYGESHGCGVGVISGCPPIPLTEADLQVLDLRRRPGQSRITSTRKET 112  
DB 2 GNTFGSLPRITTFGSHGGGVGVIIIDGCPPLRLEISPEEIQVLDLRRRPGQSKITTPRKEA 61  
QY 113 DTCKILSGTHEGVTGTPLIVPNTDQIGSHREIANVYRPSHADATYDFKYGVRVQ 172  
DB 62 DQCEILSGVFEGKLTGTPLAILVRNKDARSODYNEMAVKYPFSHADATYEAQYGINRWQ 121  
QY 173 GGRSSGRKTVGRVAGALPKKILKLGLEILSFYSKVHVVLPDADVYGSVTLQEQIES 232  
DB 122 GGRSSARETIGRVAAGAIKALQFNGVEIVAYVKSIODI---BATVDSNTVTLEQVES 178  
QY 233 NIVRCPDPEYAEKMDAIDRVRVGRDSVGGVITCVARNVPRGLGSPVDFKLESELAKAML 292  
DB 179 NIVRCPDECAEKMIERIDQVLQRKDSIGGVVECAIRNAPKGLGEPVDFKLEADLAKAM 238  
QY 293 SIPASNGEIGSGFAGTDTLQGEHNDEFYMDKAGSVRTKTRNSGGVQGSINVEIVHPKV 352  
DB 239 SLPATKGFEGSGFAGTLLTQSHNDEYLDAGEWRTRTNRSQGGVQGSINSGEPIIMRI 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVSVESMAALVLMQDMAHVAQ 412  
DB 299 AFKPTATIGQEQKTSNIGESTTLLAARGHDPVLPRAVPVVEAALVLCDDHLARFQAQ 358  
QY 413 CE 414  
DB 359 CK 360  
RESULT 7  
AROC ANASP  
ID AROC ANASP STANDARD; PRT; 362 AA.  
AC Q8YFP9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate  
DE phospholase).  
GN Name=aroC; OrderedLocusNames=all0797;  
OS Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Katanabe T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriquchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohza M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium *Anabaena* sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =  
CC chorismate + phosphate.  
CC -!- COFACTOR: Reduced flavin (By similarity).  
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
CC seventh step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SIMILARITY: Belongs to the chorismate synthase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP003583; BAB72754.1; -;  
DR PIR; AC1906; AC1906.  
DR HAMAP; MF\_00300; -; 1.  
DR InterPro; IPR000453; Chorismate\_synth.  
DR Pfam; PF01264; Chorismate\_synth; 1.  
DR ProDom; PD002941; Chorismate\_synth; 1.  
DR TIGRFAMs; TIGR00033; aroC; 1.  
DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.  
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.  
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.  
DR Aromatic amino acid biosynthesis; Complete proteome; Lyase.  
SQ SEQUENCE 362 AA; 39339 MW; 4DD32588F01C22FC CRC64;  
Query Match 53.9%; Score 1186.5; DB 1; Length 362;  
Best Local Similarity 60.6%; Pred. No. 3.8e-79;  
Matches 220; Conservative 66; Mismatches 74; Indels 3; Gaps 1;  
QY 53 GNTFGNYFOVATYGESHGCGVGVISGCPPIPLTEADLQVLDLRRRPGQSRITSTRKET 112  
DB 2 GNTFGHLPRITTFGSHGGGVGVIIIDGCPPLRLEISPEEIQVLDLRRRPGQSKITTPRKEA 61  
QY 113 DTCKILSGTHEGVTGTPLIVPNTDQIGSHREIANVYRPSHADATYDFKYGVRVQ 172  
DB 62 DTCEILSGVGEKLTGTPLISILVRNKDTRPDQYDEMAQKYPFSHADATYDAKYGINRWQ 121



Query Match 47.6%; Score 1047; DB 2; Length 364;  
 Best Local Similarity 54.9%; Pred. No. 7.2e-69;  
 Matches 200; Conservative 64; Mismatches 96; Indels 4; Gaps 2;

QY 53 GNTFGNYFOVATYGGSHGGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKET 112  
 DB 2 GSSFGKQFRVTTFGSHGGVGVILDCPPKLNIDLIQNELDRRRPQOSKITTPRNEB 61  
 QY 113 DTCKILSGTHGVTTCPIILVIVPNTDQIGSHREIANVVRPASHADATYFKYGVRAVOG 172  
 DB 62 DKLEILSLGKEGTLGTPIAMVRNKDQRPDYSNLEQVFRSHADGTVHLKIGIQAGSG 121  
 QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGVSVTLEQIES 232  
 DB 122 GGRASARETIGRVAAGAIKOLLKLNLFTEILLSWVKRIHDI---DSQVKNKNTLSKIDS 178  
 QY 233 NIVRCDDPYAEKMDAIDRVVRGDSGVGVTCTVARNVPRGLSPVFKLSESLAKAML 292  
 DB 179 NIVRCDDPYAEKMDAIDRVVRGDSGVGVTCTVARNVPRGLSPVFKLSESLAKAML 238  
 QY 293 SIPASNGFEIGSGFAGTDLTSGSHNDDEFY-MDKAGSVTRTRNRSQGVGQGISNVEIVHFK 351  
 DB 239 SIPATKGFEGISGFLGTLYRGSHNDSTFESDIIKLTISNNSGGIQGINSNGENIEMK 298  
 QY 352 VAFKPTPSIGVKQNTVSRRERQNVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVA 411  
 DB 299 IAFKPTATIGKEQKTVNADGKVKLAKGRHDPVLPRAVPVDSVVALVLDHLLHQA 358  
 QY 412 QCEN 415  
 DB 359 QCSL 362

## RESULT 10

QYU9FO PRELIMINARY; PRT; 363 AA.

ID QYU9FO  
 AC QYU9FO  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-WAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Chorismate synthase (EC 4.2.3.5).  
 GN Name=aroC; OrderedLocusNames=SYNW0308;  
 OS Synchococcus sp. (strain WH8102).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.  
 OX NCBI\_TaxID=84588;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;  
 RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,  
 Chain P., Lamerdin J.E., Regala W., Allen E.B., McCarren J.,  
 Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;  
 RT "The genome of a motile marine Synchococcus."  
 RC Nature 424:1037-1042(2003).  
 RL -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl)-3-phosphoshikimate =  
 chorismate + phosphate.  
 CC -!- COFACTOR: Reduced flavin (By similarity).  
 CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC seventh step.  
 CC -!- SIMILARITY: Belongs to the chorismate synthase family.  
 CC EMBL; BX569689; CAB06823.1; -.  
 DR HSSP; P28777; 1R52.  
 DR GO; GO:0004107; F:chorismate synthase activity; IEA.  
 DR GO; GO:0016029; F:lyase activity; IEA.  
 DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.  
 DR InterPro; IPR000453; Chorismate synth.  
 DR Pfam; PF01264; Chorismate synth; 1.  
 DR ProDom; PD002941; Chorismate synth; 1.  
 DR TIGRFAMs; TIGR00033; aroC; 1.  
 DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.  
 DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.  
 DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.  
 KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;  
 Complete proteome; Lyase.

SQ SEQUENCE 363 AA; 38266 MW; CDP87180AF34BF70 CRC64;  
 Query Match 47.2%; Score 1038; DB 2; Length 363;  
 Best Local Similarity 54.8%; Pred. No. 3.3e-68;  
 Matches 200; Conservative 65; Mismatches 96; Indels 4; Gaps 2;

QY 53 GNTFGNYFOVATYGGSHGGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKET 112  
 DB 2 GSSFGDLFRISTFGSHGGVGVIVGCPPLDLDDVDAIQAELEDRRRPQOSHITTPRKEA 61  
 QY 113 DTCKILSGTHGVTTCPIILVIVPNTDQIGSHREIANVVRPASHADATYFKYGVRAVO 171  
 DB 62 DQVELSLGLLDGQTTLTGTPIAMVRNKDQRPDYSNLEQVFRSHADATYQVKYGVQARS 121  
 QY 172 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGVSVTLEQIE 231  
 DB 122 GGRASARETIGRVAAGAIKOLLKLNLFTEILLSWVKRIHTI---EAVVDVQAVTLDAIE 178  
 QY 232 NIVRCDDPYAEKMDAIDRVVRGDSGVGVTCTVARNVPRGLSPVFKLSESLAKAM 291  
 DB 179 NIVRCDDPYAEKMDAIDRVVRGDSGVGVTCTVARNVPRGLSPVFKLSESLAKAM 238  
 QY 292 SIPASNGFEIGSGFAGTDLTSGSHNDDEFYMDKAGSVTRTRNRSQGVGQGISNVEIVHFK 351  
 DB 239 MSLPATKGFEGISGFLGTLYRGSHNDFAVPTDDRLQTATNNSGGIQGINSNGSPVIR 298  
 QY 352 VAFKPTPSIGVKQNTVSRRERQNVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVA 411  
 DB 299 VAFKPTATIRKEQKTVNADGKVKLAKGRHDPVLPRAVPVDSVVALVLDHLLHQA 358  
 QY 412 QCEN 416  
 DB 359 QCSLW 363

## RESULT 11

AROC PROMA  
 ID AROC PROMA STANDARD; PRT; 362 AA.  
 AC P46894;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate  
 phospholigase).  
 DE phospholigase).  
 GN Name=aroC; OrderedLocusNames=Pro0253;  
 OS Prochlorococcus marinus.  
 OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=1219;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SARG / CCM 1375 / SS120;  
 RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;  
 RA Dufresne A., Sallanoubat M., Partensky F., Artiguenave F., Axmann I.M.,  
 Barbe V., Duprat S., Galperin M.V., Koonin E.V., Le Gall F.,  
 Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,  
 Scanlan D.J., Tandeau de Marsac N., Weissbach J., Wincker P.,  
 Wolf Y.I., Hess W.R.;  
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,  
 a nearly minimal oxypotrophic genome."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).  
 RN [2]  
 RP SEQUENCE OF 1-52 FROM N.A.  
 RC STRAIN=SARG / CCM 1375 / SS120;  
 RX MEDLINE=95284368; PubMed=776900;  
 RA Hess W.R., Weihe A., Loiseux-De Goer S., Partensky F., Vault D.;  
 RT "Characterization of the single psbA gene of Prochlorococcus marinus  
 CCM 1375 (Prochlorophyta)."  
 RL Plant Mol. Biol. 27:1189-1196(1995).  
 CC -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl)-3-phosphoshikimate =  
 chorismate + phosphate.  
 CC -!- COFACTOR: Reduced flavin (By similarity).  
 CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;



```
CC seventh step.
CC -!- SURUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE017161; AAP99299.1; -.
DR EMBL; Z49201; CAA89063.1; -.
DR HAMAP; MF 00300; -, 1
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE_SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Lyase.
FT CONFLICT 49 52 PQGS -> RQOR (in Ref. 2).
SQ SEQUENCE 362 AA; 39429 MW; 4D761D46A694506C CRC64;

Query Match 46.5%; Score 1023.5; DB 1; Length 362;
Best Local Similarity 52.5%; Pred. No. 3.8e-67;
Matches 191; Conservative 69; Mismatches 101; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVISCPPRIPLTEADLQVELDRRRPGQSRTITRKET 112
DB 2 GSSFGDLFRVSTFGESHGSGVGIVGECPPRLDLEKIQEELDRRRPGQSKITPRKEL 61

QY 113 DTKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDPKYGVRAVQG 172
DB 62 DQVEILSGIANRETLTGTPIAMVNRDQRPDYKEMKIFRPSHADGTYHLKYGQAPSG 121

QY 173 GGRSSGRKTGVRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLEQIES 232
DB 172 GGRASARETIGRVAAGAIAKQLLQKQVNTQILAWVKRIHNI---EAEIDINTIGFADIES 178

QY 233 NIVRCPDPEYAEKIMDAIDRVVRVRSVGGVITCVARNVPRGLSPVPDKLESELAKAML 292
DB 179 NIVRCPNQDVAKLMQRIEETSRDSCGGLIECVRVNVPAGLGNPVPDKLEADLSKALM 238

QY 293 SIPASNGFEIGSGFAGTDLTSGHNDEFYMDKAGSVRTRTRNSGGVQGGISNVEIVHFKV 352
DB 239 SLPATKGPEVSGFRGTFTLKGSEHNDAFIAGDKNRLRTATNNSGGIQGGISNGSPILRV 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQMLMAHVAQ 412
DB 299 GFKPATIRKQQQTIDSEKQITLASKGRHDPVCLPRAVPVMEAMVSVTLADHLLRQRGQ 358

QY 413 CEMF 416
DB 359 CSLW 362

RESULT 12
Q7V4Y9 PRELIMINARY; PRT; 362 AA.
AC Q7V4Y9;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Chorismate synthase (EC 4.2.3.4).
GN Name=aroC; OrderedLocusNames=PMT1795;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
CC -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl) -3-phosphoshikimate =
CC chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
DR EMBL; BX572100; CAE21970.1; -.
DR HSSP; P28777; IR52.
DR GO; GO:0003856; F:3-dehydroquinase synthase activity; IEA.
DR GO; GO:0004107; F:chorismate synthase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE_SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
KW Complete proteome; Lyase.
SQ SEQUENCE 362 AA; 36714 MW; 435DEA6777BB86D3 CRC64;

Query Match 45.8%; Score 1007.5; DB 2; Length 362;
Best Local Similarity 51.6%; Pred. No. 5.8e-66;
Matches 188; Conservative 69; Mismatches 104; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVISCPPRIPLTEADLQVELDRRRPGQSRTITRKET 112
DB 2 GSSFGDLFRVSTFGESHGSGVGIVGECPPRLDLEKIQEELDRRRPGQSKISTPRKEE 61

QY 113 DTKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDPKYGVRAVQG 172
DB 62 DQVEILSGIANRETLTGTPIAMVNRDQRPDYKEMKIFRPSHADATYQAKYGIQARSQ 121

QY 173 GGRSSGRKTGVRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLEQIES 232
DB 172 GGRASARETIGRVAAGAIAKQLLQKQVNTQILAWVKRIHNI---EAEINAQDVSIDVEA 178

QY 233 NIVRCPDPEYAEKIMDAIDRVVRVRSVGGVITCVARNVPRGLSPVPDKLESELAKAML 292
DB 179 NIVRCPNQDVAAQVVERIEAISREGDSCGVIECVRVNAPMGLOMPVPDKLEADLAKAVM 238

QY 293 SIPASNGFEIGSGFAGTDLTSGHNDEFYMDKAGSVRTRTRNSGGVQGGISNVEIVHFKV 352
DB 239 SLPATKGFEIGSGFGTLLKGSEHNDALFSPNDGLRTATNNSGGIQGGITNGESIVRV 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQMLMAHVAQ 412
DB 299 AFKPTATIRKQQQTIDAGDGNNTTILSAKGRHDPVCLPRAVPVMEAMVSVTLADHLLRQOQG 358

QY 413 CEMF 416
DB 359 CSLW 362

RESULT 13
AROC SCHPO STANDARD; PRT; 395 AA.
ID _AROC SCHPO
AC 074413; Q9Y7Q4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
```



Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase).

ORFNames=SPCC1223.14, SPCC297.01;

OS Schizosaccharomyces pombe (Piscesion yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

NCBI\_TaxID=4896;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=972;

RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;

RX Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Stevens S., Warren T., Whitehead S., Taylor K.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkheart G., Aert R., Robben J., Grymonprez B., Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RA "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate = chorismate + phosphate.

CC -!- COFACTOR: Reduced flavin (By similarity).

CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; seventh step.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SIMILARITY: Belongs to the chorismate synthase family.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by not-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; AL031579; CAA20883.1; --

DR EMBL; AL049609; CAB40781.1; --

DR FIR; T41268; T41268

DR GeneDB\_Spombe; SPCC1223.14; --

DR InterPro; IPR000453; Chorismate synth.

DR Pfam; PF01264; Chorismate synth; 1.

DR ProDom; PD002941; Chorismate synth; 1.

DR TIGRFAMs; TIGR00033; arcc; 1

DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.

DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.

DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.

KW Aromatic amino acid biosynthesis; lyase.

SQ SEQUENCE 395 AA; 42325 MW; B551505160696DE5 CRC64;

Query Match 43.2%; Score 951; DB 1; Length 395;

Best Local Similarity 51.3%; Pred. No. 9.4e-62;

Matches 194; Conservative 57; Mismatches 103; Indels 24; Gaps 4;

Db 2 SSFGTLFKVTTVGESHCKSVGCVGCPGNNLTESDVQVLTTRRRPGQSNLTTPRNEKD 61

QY 114 TKILSGTHGVTGTPILVVPNTDQIGSDHREANVVRSHADATYDFKYGVRVQGG 173

Db 62 KVQIQSGTFEGVTGTPILVVPNTDQIGSDHREANVVRSHADATYDFKYGVRVQGG 121

QY 174 GRSSGRKTVGRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED--- 218

Db 122 GSSARETIGRVAAGIAEKYLLLEAYGVVAVFVSUGKIALPLHETASSAILDEDDTF 181

QY 219 ----AVDY-----GSVLEQIESNIVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARN 270

Db 182 ESPITAEYLKFLNKLTREEDVKTTRVCPHAATAAKMAERITRARNHDSIGGTVTCVIRN 241

QY 271 VPRGIGSPVFDKLESELAKAMLSIPASNGFEGISGAGTDLTGSEHDEFFYMD-KAGSVR 329

Db 242 VPTGLGEFCFDKLEAKLAHAMWSIPATKSPISGREGCKVAGSKHNDLFYRNADTGKLG 301

QY 330 TRTRNSGGVQGGISNVEIVHFVKAPKTPSPISGKONTVSRERQNVELLARGHDPVAPR 389

Db 302 TLTNNSGGVQGGISNGENVYTFGKSPATIGVEQSTSRYDGSQDGLAAGKGRHDFCVPR 361

QY 390 AVPVVSEMAALVMDQLM 407

Db 362 AIPIVEAMALVMDAVM 379

RESULT 14

Q6AIP3 PRELIMINARY; PRT; 382 AA.

ID Q6AIP3

AC Q6AIP3

DT 25-OCT-2004 (TREMREL. 28, Created)

DT 25-OCT-2004 (TREMREL. 28, Last sequence update)

DE Probable chorismate synthase.

GN OrderedLocustNames=DP3058;

OS Desulfotalea psychrophila.

OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;

OC Desulfobubaceae; Desulfotalea.

OX NCBI\_TaxID=84980;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LSV54 / DSM 12343;

RX PubMed=15305914;

RA Rabus R., Ruepp A., Frickey T., Rattei T., Partmann B., Stark M., Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K., Teeling H., Leuschner W.D., Gloeckner P.-O., Lupas A.N., Amann R., Klenk H.-P.;

RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium from permanently cold Arctic sediments.";

RL Environ. Microbiol. 6:887-902(2004).

CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate = chorismate + phosphate.

CC -!- COFACTOR: Reduced flavin (By similarity).

CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; seventh step.

CC -!- SIMILARITY: Belongs to the chorismate synthase family.

DR EMBL; CR522870; CAG37787.1; --

DR GO; GO:0004107; P:chorismate synthase activity; IEA.

DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.

DR InterPro; IPR000453; Chorismate synth.

DR Pfam; PF01264; Chorismate synth; 1.

DR ProDom; PD002941; Chorismate synth; 1.

DR TIGRFAMs; TIGR00033; arcc; 1.

DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.

DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.

DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.

KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;

KW Complete proteome; lyase.

SQ SEQUENCE 382 AA; 41321 MW; 5A93D8AF21802F4A CRC64;

Query Match 43.2%; Score 950.5; DB 2; Length 382;

```
Best Local Similarity 51.8%; Pred. No. 9.8e-62;
Matches 188; Conservative 65; Mismatches 107; Indels 3; Gaps 2;

Qy 54 NTFGNFYQVATYGESHGGVGVCGVISCPPRIPLTEADLQVELDRRRPQGSRIITSTRKETD 113
Db 23 SSFGTLXKYSTYGESHCKGKGVVWIDGCPGLESEADIQQLDRRRPQGVNLTTRKEAD 82
Qy 114 TCKILSGTHEGVTGTPILVIVPNTDOIGSDHREIANVVRSHADATYDFKYGVRVQGG 173
Db 83 QVMILSGTENGKTLGTPIALHVNRRDQRPDGYDGLPRPSHADFTYQMYKIGRASGG 142
Qy 174 GRSSGRKTVGRVAAGALPKKILKCGLEILSFVSKVHVLPEDAVDYGSVTLLEQIESN 233
Db 143 GRASARETIGTVAAGAVKVLKXKYGIEIVAVESVGP1--QASGINPDTISREVDKQ 200
Qy 234 IVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLSGVPFKLSELAAMLS 293
Db 201 ISRCPDQITAEEMKFPITELKBERGSTGIVTCCIRNSPVLGEPYIEKLEAKLAQAMLA 260
Qy 294 IPASNGFIEGSGFAGTDLTGSEHNDREYMDKAGSVRTNRSVGGVQGGISNVEIVHFKVA 353
Db 261 IPASNGFIEGSGFAGSRLGSEHNDLFIW-KDNRLGTSTNNSGGIQQGSGISNGEPILFRVA 319
Qy 354 FKPTPSIGVKQNTVSRERQNVELLARGHRDPCVAPRAVPVSVESMAALVMDQMAHVAQC 413
Db 320 FAPTATISLPQETATFTGKTETMQAKGRHDPCCVSRAP1VSVESMAALVLLDMWRQESRK 379
Qy 414 EMP 416
Db 380 SPF 382

RESULT 15
Q6CNY1 PRELIMINARY; PRT; 376 AA.
AC Q6CNY1;
DC 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome E of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN ORFNames=KLI1A0E09031G;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boirame A., Boyer J., Catholico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekaja P., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weissenbach J.,
RA Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
chorismate + phosphate.
CC
```

```
CC -1- COFACTOR: Reduced flavin (By similarity).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
seventh step.
CC -1- SIMILARITY: Belongs to the chorismate synthase family.
DR EMBL; CR382125; CAG99445.1; -.
DR GO; GO:0004107; F:chorismate synthase activity; IEA.
DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01284; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroc; 1.
DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis; Lyase.
SQ SEQUENCE 376 AA; 40348 MW; D4CEB8B2B4DF170A9 CRC64;

Query Match 42.9%; Score 944; DB 2; Length 376;
Best Local Similarity 51.9%; Pred. No. 2.9e-61;
Matches 188; Conservative 62; Mismatches 104; Indels 8; Gaps 3;

Qy 54 NTFGNFYQVATYGESHGGVGVCGVISCPPRIPLTEADLQVELDRRRPQGSRIITSTRKETD 113
Db 2 STFGQIFRVATYGESHCCKSVGCIVDGVPFGMSLTDDIQQLTRRRPQSKLSTPRNEKD 61
Qy 114 TCKILSGTHEGVTGTPILVIVPNTDOIGSDHREIANVVRSHADATYDFKYGVRVQGG 173
Db 62 RVEIQSGTEBFGKLTGTPIGMLVKNEDQRPHDYSMDNYPSPSHADFTYSSKYGIKASSGG 121
Qy 174 GRSSGRKTVGRVAAGALPKKILKCGLEILSFVSKVHVLPEDAVD-----YGSVTL 227
Db 122 GRSSARETIGRVAAGAIAEKFLQLSNVEIVAFVTOIGAVKMRDPODPKFEQVNLNITR 181
Qy 228 EQIE-SNIVRCPPDEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLSGVPFKLSE 286
Db 182 EKVDAGPIRCPDSSVADAMVKIEIERYRGEKDSIGGVVTCVIRNLPAGLGEPCPKLEAL 241
Qy 287 LAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDREYMDK-AGSVTRTRNRSVGGVGGISNV 345
Db 242 LAHAWLSIPASKGFEIGSGFGDVCCLPGSKHNDPFYFDEDTGRUKTKTNNSGGIQQGINSNG 301
Qy 346 EIVHFKVAFKPTPSIGVKQNTVSRERQNVELLARGHRDPCVAPRAVPVSVESMAALVMDQ 405
Db 302 ENIYFSVPFKSAATISQEQATATYDGKSGILAAKGRHDPSTVTPRATPIVEAMAAVLADA 361
Qy 406 LM 407
Db 362 VL 363

Search completed: August 25, 2005, 19:31:17
Job time : 71 secs
```

GenCore version 5.1.6  
: Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:28:05 ; Search time 44 Seconds  
(without alignments)  
727.828 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKPOQVAHSRARLAPR.....VAQCEPALTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

'Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	429	4	US-09-743-207-8
2	1797.5	81.7	440	4	US-09-743-207-2
3	1541	70.0	431	4	US-09-103-331-39
4	1541	70.0	431	4	US-09-631-594-48
5	1257	57.1	307	4	US-09-743-207-12
6	1191.5	54.2	362	4	US-09-103-331-38
7	1191.5	54.2	362	4	US-09-631-594-47
8	1066	48.5	257	4	US-09-743-207-6
9	938	42.6	376	4	US-09-103-331-42
10	938	42.6	376	4	US-09-631-594-51
11	930	42.3	432	4	US-09-103-331-40
12	930	42.3	432	4	US-09-631-594-49
13	859	39.0	377	4	US-09-489-039A-12546
14	840.5	38.6	362	4	US-09-543-681A-4292
15	840.5	38.2	357	4	US-09-103-331-41
16	820.5	38.2	357	4	US-09-631-594-50
17	820.5	37.3	359	4	US-09-248-796A-18202
18	786	35.7	379	4	US-09-252-991A-30159
19	741	33.7	367	4	US-09-328-352-7662
20	731	33.2	536	4	US-09-103-331-2
21	729	33.1	374	4	US-09-540-236-2230
22	727	33.0	536	4	US-09-631-594-46
23	644	29.3	527	4	US-09-103-331-4
24	644	29.3	527	4	US-09-631-594-55
25	642.5	29.2	388	4	US-09-438-185A-1038
26	641	29.1	376	4	US-09-198-452A-1112
27	611	27.8	339	4	US-09-902-540-16343

28	478.5	21.8	168	4	US-09-743-207-14	Sequence 14, Appl
29	459.5	20.9	392	4	US-09-107-532A-4840	Sequence 4840, Ap
30	437	19.9	425	3	US-09-134-001C-3143	Sequence 3143, Ap
31	419.5	19.1	391	4	US-09-107-433-4630	Sequence 4630, Ap
32	416.5	18.9	388	4	US-09-583-110-3543	Sequence 3543, Ap
33	413.5	18.8	395	4	US-09-134-000C-4982	Sequence 4982, Ap
34	380.5	17.3	353	4	US-09-710-279-2040	Sequence 2040, Ap
35	353	16.0	108	4	US-09-743-207-4	Sequence 4, Appl
36	295	13.4	81	4	US-09-743-207-10	Sequence 10, Appl
37	102.5	4.7	432	4	US-09-252-991A-22020	Sequence 22020, A
38	101	4.6	96	4	US-09-248-796A-18203	Sequence 18203, A
39	97.5	4.4	434	4	US-09-252-991A-27401	Sequence 27401, A
40	97	4.4	233	4	US-09-328-352-5981	Sequence 5981, Ap
41	96.5	4.4	402	4	US-09-252-991A-27689	Sequence 27689, A
42	96.5	4.4	711	4	US-09-252-991A-22259	Sequence 22259, A
43	96	4.4	304	4	US-09-489-039A-13414	Sequence 13414, A
44	96	4.4	407	4	US-09-252-991A-21346	Sequence 21346, A
45	96	4.4	409	4	US-09-252-991A-21338	Sequence 21338, A

## ALIGNMENTS

### RESULT 1

US-09-743-207-8  
; Sequence 8, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: US/09/743,207  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-743-207-8

Query Match	100.0%	Score 2200;	DB 4;	Length 429;
Best Local Similarity	100.0%	Pred. No. 1.8e-234;		
Matches 429;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MTTVPKPOQVAHSRARLAPRAIGALLEPAPASSLSRFAVHRCRTARLEVKASGNTFGNYF	60	
Db	1	MTTVPKPOQVAHSRARLAPRAIGALLEPAPASSLSRFAVHRCRTARLEVKASGNTFGNYF	60	
Qy	61	QVATYVSHGGVGVISGCPPIPLTEADLOVELDRRPGOSRITSTRKETDCKILSG	120	
Db	61	QVATYVSHGGVGVISGCPPIPLTEADLOVELDRRPGOSRITSTRKETDCKILSG	120	
Qy	121	THEGVTTGTPILVTPNTDQIGSDHREIANYRPSHADATYDFKYGVRAVQGGRSSGRK	180	
Db	121	THEGVTTGTPILVTPNTDQIGSDHREIANYRPSHADATYDFKYGVRAVQGGRSSGRK	180	
Qy	181	TVGRVAAGALKPKILKLCGLEILSFVSKVHQVLPEDAVDYGVSVTLEQIESNIVRCDDP	240	
Db	181	TVGRVAAGALKPKILKLCGLEILSFVSKVHQVLPEDAVDYGVSVTLEQIESNIVRCDDP	240	
Qy	241	EYAEKMDAIDRVVRGDSVCGVITCVARNVPRGIGSPVDFKLESELAKAMLSIPASNGF	300	
Db	241	EYAEKMDAIDRVVRGDSVCGVITCVARNVPRGIGSPVDFKLESELAKAMLSIPASNGF	300	
Qy	301	EIGSGFACTDLTGTSEHNDDEFYMDKAGSVRTRTNSGGVGGISNVEIVHFVKVAPKPTPSI	360	

Db 301 EIGSGFAGTDLTSGSEHNDYFMDKAGSVKTRTRNRSQGGVGGISNVEIVHFVKVAFKPTPSI 360  
Qy 361 GVKONTVSRERONVELLARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFALNT 420  
Db 361 GVKONTVSRERONVELLARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFALNT 420  
Qy 421 ALQEPVGSF 429  
Db 421 ALQEPVGSF 429  
  
RESULT 2  
US-09-743-207-2  
; Sequence 2, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 440  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-743-207-2

Query Match 81.7%; Score 1797.5; DB 4; Length 440;  
Best Local Similarity 81.9%; Pred. No. 6.8e-190;  
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;  
  
Qy 4 VPKPOQVAHSRARLAPRAIGALLFPAPASSLSLRFV-HRCRTARLEVKASGNTFGNYFQV 62  
Db 5 VSQPPVSAASRFLPRGIGALPESAP--TSLSVGRRRRASSLLEVKASGNVFGNYFQV 62  
  
Qy 63 ATYGESHGGVGCVISGCPPRIPLTEADLOVELDRRRPQSRITSTRKETDTCKILSGTH 122  
Db 63 ATYGESHGGVGCVISGCPPRIPLTEADMQVELDRRRPQSRITSTRKETDTCKILSGTH 122  
  
Qy 123 EGVTTGTPILVIPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQGGRRSGRKT 182  
Db 123 DGMVTGTPIHVPNTDQGGDYSEMSKAYRPSHADATYDFKYGVRAVQGGRRSARETI 182  
  
Qy 183 GRVAAGALPKILKUKLKGLEILSFVSKVHQVVLPEDAVDYGSVTLQIESNIVRCPDPEY 242  
Db 183 GRVAAGALAKTILKLSGVEILAFYSKVHQVVLPEDAVDYETVTLTEHTESNIVRCPDPEY 242  
  
Qy 243 AEKMTDAIDRVVRGDSVGGVITCVARNVPRGLSGSPVDKLESELAKWLSIPASNGPEI 302  
Db 243 AEKMTAAIDTVRVRGDSVGGVITCVARNVPRGLSGSPVDKLESELAKWLSIPASKGPEI 302  
  
Qy 303 GSGFAGTDLTSGSEHNDYFMDKAGSVKTRTRNRSQGGVGGISNVEIVHFVKVAFKPTPSIGV 362  
Db 303 GSGFAGTDLTSGSEHNDYFMDKAGSVKTRTRNRSQGGVGGISNVEIVHFVKVAFKPTPSIGV 362  
  
Qy 363 KONTVSRERONVELLARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFALNTAL 422  
Db 363 KONTVSRERONVELLARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFALNTAL 422  
  
Qy 423 QEPVGS 428  
Db 423 QEPVGS 428

RESULT 3

US-09-103-331-39  
; Sequence 39, Application US/09103331  
; Patent No. 6693654  
; GENERAL INFORMATION:  
; APPLICANT: McLeod, Rima W.  
; APPLICANT: Roberts, Craig W.  
; APPLICANT: Roberts, Fiona  
; APPLICANT: Johnson, Jennifer J.  
; APPLICANT: Mets, Laurens J.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC  
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN  
; TITLE OF INVENTION: PARASITE COMPONENTS  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/103,331  
; APPLICATION NUMBER: US/09/103,331  
; FILING DATE: 23-JUN-1998  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Martin, Alice O.  
; REGISTRATION NUMBER: 35,601  
; REFERENCE/DOCKET NUMBER: 8336/9  
; TELEPHONE: 312-321-4200  
; TELEFAX: 312-321-4299  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 431 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-103-331-39  
  
Query Match 70.0%; Score 1541; DB 4; Length 431;  
Best Local Similarity 72.7%; Pred. No. 1.6e-161;  
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;  
  
Qy 32 SSSLRFVHRCRTARLEVKASGNTFGNYFOVATYGESHGCGVGVISGCPPRIPLTEADL 91  
Db 29 SSSLRFPTHRSQPKRLIEQAAGNTFGNYFRVTFYGESHGCGVGVICIDGCPPRLPLSMD 88  
  
Qy 92 QVELDRRRPQSRITSTRKETDTCKILSGTHGVTTCPIILVIVPNTDQIGSDHREIANV 151  
Db 89 QVELDRRRPQSRITSTRKETDTCKILSGTHGVTTCPIILVIVPNTDQIGSDHREIANV 148  
  
Qy 152 YRPSHADATYDFKYGVRAVQGGRRSGRKTGRVRAAGALPKILKKGLEILSFVSKVH 211  
Db 149 YRPSHADATYDFKYGVRAVQGGRRSGRKTGRVRAAGALPKILKKGLEILSFVSKVH 208  
  
Qy 212 QVVLPEDAVDYGSVTLQIESNIVRCPDPEYAEKMTDAIDRVVRGDSVGGVITCVARNV 271  
Db 209 NVVLPEDLVNQIVTLQIESNIVRCPNPEYAEKMTDAIDRVVRGDSVGGVITCVARNV 268  
  
Qy 272 PRGLSGSPVDKLESELAKWLSIPASNGFETIGSGFAGTDLTSGSEHNDYFMDKAGSVKTR 331  
Db 269 PRGLGTPVFDKLEELAKWLSIPATKGFEGSGFAGTDLTSGSEHNDYFMDKAGSVKTR 328  
  
Qy 332 TNRSGGVQGGISNVEIVHFVKVAFKPTPSIGVKONTVSRERONVELLARGHRDPCVAPRAV 391  
Db 329 TNRSGGIQGGISNVEIVHFVKVAFKPTPSIGVKONTVSRERONVELLARGHRDPCVAPRAV 388

QY 392 PVVSMALVLDQMAHVAQCEMFALNTALQEPV 426  
Db 389 PMVEAMVALVLDQMTQYAQCMLFPVNLTLQEP 423

## RESULT 4

US-09-631-594-48  
; Sequence 48, Application US/09631594  
; Patent No. 6737237

## GENERAL INFORMATION:

; APPLICANT: MCLEOD, RIMA W.  
; APPLICANT: ROBERTS, CRAIG W.  
; APPLICANT: ROBERTS, FIONA  
; APPLICANT: JOHNSON, JENNIFER J.  
; APPLICANT: KIRLSITS, MICHAEL  
; APPLICANT: FERGUSON, DAVID  
; APPLICANT: LYONS, RUSSELL  
; APPLICANT: MUI, ERNEST  
; APPLICANT: MACK, DOUG  
; APPLICANT: SAMUEL, BENJAMIN  
; APPLICANT: GORNICKI, PIOTR  
; APPLICANT: ZUTHER, ELEN

; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES  
; FILE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS

; CURRENT APPLICATION NUMBER: 19338-90966

; FILE REFERENCE: 19338-90966

; CURRENT FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: PCT/US00/11478

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: PCT/US97/12497

; PRIOR FILING DATE: 1997-07-18

; NUMBER OF SEQ ID NOS: 83

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 48

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Solanum lycopersicum

US-09-631-594-48

Query Match 70.0%; Score 1541; DB 4; Length 431;

Best Local Similarity 72.7%; Pred. No. 1.6e-161;

Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 32 SSSLFVHRCRTARLEKVASGNTGNYFOVATYGESGGVGVCGVSGCPPIPLTEADL 91  
Db 29 SSSLFPPTTHRSQPKLEIQAGNTFGNYPRVTTFGESGGVGGCIIIDGCPPLPLUSEDM 88

QY 92 QVLEDRRPGQSRITSTRKETDTCKILSGTHGVTGTPIVIVPNTDQIGSDHREIANV 151  
Db 89 QVLEDRRPGQSRITTPKRTDTCKISGTADGLTGSPIKVEVNTDORNDYSEMSLA 148

QY 152 YRPSHADATYDFKYGVRAVQGGRRSGRTKTVGRVAAGALPKILKLCGLEILSFVSKVH 211  
Db 149 YRPSHADATYDFKYGVRSVQGGRRSARETIGRAVAGAVAKILKLYSGTEILAVSQVH 208

QY 212 QVLEPDAVDYGSVTLQIESNIVRCPPDEYAEKMDAIDRVVRGDSVGGVITCVARNV 271  
Db 209 NVVLPEDLVNDQIVTLQIESNIVRCNPENYAEKMGIDYVVRGDSVGGVITCVARNV 268

QY 272 PRGLSPVDFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDFFYMDKAGSVTR 331  
Db 269 PRGLTVPVDFDKLESELAKAMLSIPATKGFEGSGFAGTGMTGSEHNDFFYMDHDIQTK 328

QY 332 TNRSGVQGGISNVEIVHFKVAFKPTPSIGVQNTVSRERQNVLLARGHDPVAPRAV 391  
Db 329 TNRSGIQGGISNVEIVHFKVAFKPTPSIARQNTVSRERQNVLLARGHDPVAPRAV 388

QY 392 PVVSMALVLDQMAHVAQCEMFALNTALQEPV 426  
Db 389 PMVEAMVALVLDQMTQYAQCMLFPVNLTLQEP 423

## RESULT 5

US-09-743-207-12  
; Sequence 12, Application US/09743207  
; Patent No. 6653531

## GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-743-207-12

Query Match 57.1%; Score 1257; DB 4; Length 307;

Best Local Similarity 81.9%; Pred. No. 2.6e-130;

Matches 245; Conservative 22; Mismatches 26; Indels 6; Gaps 2;

QY 15 ARLAPRAIG----ALLEPAPASSLSRPAVHRCRTARLEKVASGNTFGNYFOVATYGESHG 70  
Db 11 ARVLPRGGGGGFRAPPESAPA--SLRFSVGRRAARLEKVASANVFGNYFOVATYGESHG 68

QY 71 GGVCVLSGCPPIPLTEADLQVLEDRRPGQSRITSTRKETDTCKILSGTHGVTGTCTP 130  
Db 69 GGVCVLSGCPPIPLTEADLQVLEDRRPGQSRITSTRKETDTCKILSGTHGVTGTCTP 128

QY 131 ILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQGGRRSGRTKTVGRVAAGAL 190  
Db 129 IHVFPVNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQGGRRSGRTKTVGRVAAGAL 188

QY 191 PKKILKLCGLEILSFVSKVHQQVLPDAVDYGSVTLQIESNIVRCPPDEYAEKMDAI 250  
Db 189 AKKILKLCGLEILSFVSKVHQQVLPDAVDYGSVTLQIESNIVRCPPDEYAEKMDAI 248

QY 251 DRVVRGDSVGGVITCVARNVPRGDSVDFDKLESELAKAMLSIPASNGFEIGSGFACT 309  
Db 249 DKVRVGDIGGVVTCIARNVPRGDSVDFDKLESELAKAMLSIPASNGFEIGSGFACT 307

## RESULT 6

US-09-103-331-38  
; Sequence 38, Application US/09103331  
; Patent No. 6699654

## GENERAL INFORMATION:

; APPLICANT: McLeod, Rima W.  
; APPLICANT: Roberts, Craig W.  
; APPLICANT: Roberts, Fiona  
; APPLICANT: Johnson, Jennifer J.  
; APPLICANT: Metz, Laurens J.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC  
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN  
; TITLE OF INVENTION: PARASITE COMPONENTS  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; STREET: Plaza Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103.331
; FILING DATE: 23-JUN-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8336/9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-103-331-38

Query Match          54.2%; Score 1191.5; DB 4; Length 362;
Best Local Similarity 62.2%; Pred. No. 6.1e-123;
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYQVATYGESHGCGVCGVSCPPRIPLTEADLOVELDRRRPGQSRTSTRKET 112
Db 2 GNTFGSLFRITTRGESHGGGVIIIDGCPRLRLEISPEEIQVLDLDRRRPGQSRTSTRKET 61
QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVAVQ 172
Db 62 DQCEILSGVFEGKTLGTPAILVRNKDARSQDYNEMAVKYPESHADATYEAKEYGIRNWQ 121
QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIES 232
Db 122 GGRSSARETIGRVAAGAIAKKILAQFNGVEIVAVYKSIQDI---EATVDSNTVTLEQVES 178
QY 233 NIVRCPDEYAEKMDAIDRVVRVGDVSGVITCVARNVPRGLSGSPVDPKLESELAKAML 292
Db 179 NIVRCPDECAEKMIERIDQVLRQKDSIGGVVECAIRNAPKGLGEPVDFDKLEADLAKAM 238
QY 293 SIPASNGEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNSGGVQGGISNVEIVHFKV 352
Db 239 SLPATKGFEGSGFAGTLLTGSHQNDYILDEAGEWRTTRNSGGVQGGISNGEPIIMRI 298
QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQDMAHVAQ 412
Db 299 AFKPTATIGQEQKTVSNIGEETTLAAKGRHDPVLPRAVPVVEAMAALVLCDDLRLRFOAQ 358
QY 413 CE 414
Db 359 CK 360

RESULT 7
US-09-631-594-47
; Sequence 47, Application US/09631594
; Patent No. 6737237
; GENERAL INFORMATION:
; APPLICANT: MCLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, FIONA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KIRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUI, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN

;
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS
; FILE REFERENCE: 19338-90966
; CURRENT APPLICATION NUMBER: US/09/631,594
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/US00/11478
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US97/12497
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Synecchocystis sp.
; US-09-631-594-47

Query Match          54.2%; Score 1191.5; DB 4; Length 362;
Best Local Similarity 62.2%; Pred. No. 6.1e-123;
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYQVATYGESHGCGVCGVSCPPRIPLTEADLOVELDRRRPGQSRTSTRKET 112
Db 2 GNTFGSLFRITTRGESHGGGVIIIDGCPRLRLEISPEEIQVLDLDRRRPGQSRTSTRKET 61
QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVAVQ 172
Db 62 DQCEILSGVFEGKTLGTPAILVRNKDARSQDYNEMAVKYPESHADATYEAKEYGIRNWQ 121
QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLEQIES 232
Db 122 GGRSSARETIGRVAAGAIAKKILAQFNGVEIVAVYKSIQDI---EATVDSNTVTLEQVES 178
QY 233 NIVRCPDEYAEKMDAIDRVVRVGDVSGVITCVARNVPRGLSGSPVDPKLESELAKAML 292
Db 179 NIVRCPDECAEKMIERIDQVLRQKDSIGGVVECAIRNAPKGLGEPVDFDKLEADLAKAM 238
QY 293 SIPASNGEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNSGGVQGGISNVEIVHFKV 352
Db 239 SLPATKGFEGSGFAGTLLTGSHQNDYILDEAGEWRTTRNSGGVQGGISNGEPIIMRI 298
QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQDMAHVAQ 412
Db 299 AFKPTATIGQEQKTVSNIGEETTLAAKGRHDPVLPRAVPVVEAMAALVLCDDLRLRFOAQ 358
QY 413 CE 414
Db 359 CK 360

RESULT 8
US-09-743-207-6
; Sequence 6, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Triticum aestivum
; US-09-743-207-6
```

Query Match 48.5%; Score 1066; DB 4; Length 257;  
Best Local Similarity 82.7%; Pred. No. 2.7e-109;  
Matches 201; Conservative 27; Mismatches 15; Indels 0; Gaps 0;  
QY 186 AAGALPKKILKCGLEILSFVSKVHQVLPEDAVDYGVTLEQIESNTVRCPPDEYAEK 245  
DB 3 AAGAVAKILKLCGVEILAFVSKVHQVLPEDAVDYETLTLDQIESNCRCPDPEYAO 62  
QY 246 MIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFKLESEKAKMLSTPASNGEIGSG 305  
DB 63 MIDAIDKVRNGSIGGVVTCIARNVPRGLGSPVDFKLESEKAKMLSTPASNGEIGSG 122  
QY 306 FAGTDLTSGSEHDEFTYMDKAGSVRTNRSGVQGGISNVEIVHFKVAFKPTPSIGVKON 365  
DB 123 FAGTDLTSGSEHDEFTYMDKAGSVRTNRSGVQGGISNVEIVHFKVAFKPTPSIGVKON 182  
QY 366 TVSRERQNVLLIARGHDCPVAPRVVSVESMAALVMDQLMAHVAQCBMFALNTALQEP 425  
DB 183 TVTROHEDIETLLTRGRHDCPVAPRVVSVESMAALVMDQLMAHVAQCBMFALNTALQEP 242  
QY 426 VGS 428  
DB 243 IGS 245

RESULT 9  
US-09-103-331-42  
; Sequence 42, Application US/09103331  
; Patent No. 6699654  
; GENERAL INFORMATION:  
; APPLICANT: McLeod, Rima W.  
; APPLICANT: Roberts, Craig W.  
; APPLICANT: Roberts, Fiona  
; APPLICANT: Johnson, Jennifer J.  
; APPLICANT: Metz, Laurens J.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC  
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN  
; TITLE OF INVENTION: PARASITE COMPONENTS  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09103,331  
; FILING DATE: 23-JUN-1998  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Martin, Alice O.  
; REGISTRATION NUMBER: 35,601  
; REFERENCE/DOCKET NUMBER: 8336/9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-321-4200  
; TELEFAX: 312-321-4299  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-103-331-42

Query Match 42.6%; Score 938; DB 4; Length 376;  
Best Local Similarity 49.9%; Pred. No. 7.6e-95;  
Matches 183; Conservative 67; Mismatches 109; Indels 8; Gaps 3;  
QY 54 NTFNGYFOVATYGGSHGGVCGVIGCPPIPLTEADLQVELDRRRPQOSRITSTRKETD 113  
DB 2 STFGKLFRTVYGGSHGCKSVGCIVDGVPFGMSLTEADIQPOLTRRRPQOSKLTSPRDSKD 61  
QY 114 TCKILSGHEGVTGTPIVLVPTNDQIGSHREIFANVYRPSHADATYDFKYGVRAVQGG 173  
DB 62 RVETQSGTEFGTKLTPIAMMKNEQDQPHDISDMDFPRFSHADFTYSEKIGIKASGG 121  
QY 174 GRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVD-----YGSVTL 227  
DB 122 GRASARETIGRVASGAIAEKFLAQNVSVEIVAFVQIGEIKNRDSFDPBFQHLNTITR 181  
QY 228 FOIES-NIVRCPDPEYAEKMDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFKLESE 286  
DB 182 EKVDSMGPIRCPDASVAGLMTVEIKYRGNKDSIGGVVTCVVRNLPTGLGEPDFKLEAM 241  
QY 287 LAKAMLSTPASNGEIGSGFAGTDLTSGSEHDEFTYMDK-AGSVRTRTNRSGVQGGISNV 345  
DB 242 LAHAMLSTPASNGEIGSGFAGTDLTSGSEHDEFTYMDK-AGSVRTRTNRSGVQGGISNV 301  
QY 346 EIVHFKVAFKPTPSIGVKQNTVSRERQNVLLIARGHDCPVAPRVVSVESMAALVMDQ 405  
DB 302 ENIVFSVPFKSVATISQOKTATYDGBEGILAAKGRHDPATVTPRAIPVEAMTALVLADA 361  
QY 406 LMAHVAQ 412  
DB 362 LLIQKAR 368

RESULT 10  
US-09-631-594-51  
; Sequence 51, Application US/09631594  
; Patent No. 6737237  
; GENERAL INFORMATION:  
; APPLICANT: McLeod, Rima W.  
; APPLICANT: Roberts, Craig W.  
; APPLICANT: Roberts, Fiona  
; APPLICANT: Johnson, Jennifer J.  
; APPLICANT: Kirisits, Michael  
; APPLICANT: Ferguson, David  
; APPLICANT: Lyons, Russell  
; APPLICANT: Mui, Ernest  
; APPLICANT: Mack, Doug  
; APPLICANT: Samuel, Benjamin  
; APPLICANT: Gornicki, Piotr  
; APPLICANT: Zuther, Ellen  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES  
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS  
; FILE REFERENCE: 19338-90966  
; CURRENT APPLICATION NUMBER: US/09/631,594  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/US00/11478  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US97/12497  
; PRIOR FILING DATE: 1997-07-18  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-631-594-51

Query Match 42.6%; Score 938; DB 4; Length 376;  
Best Local Similarity 49.9%; Pred. No. 7.6e-95;  
Matches 183; Conservative 67; Mismatches 109; Indels 8; Gaps 3;  
QY 54 NTFNGYFOVATYGGSHGGVCGVIGCPPIPLTEADLQVELDRRRPQOSRITSTRKETD 113  
DB 2 STFGKLFRTVYGGSHGCKSVGCIVDGVPFGMSLTEADIQPOLTRRRPQOSKLTSPRDSKD 61



Db 2 STFGKLFRTVTTGESHCKSVGCI VDGVPVPGMSL TEADIQPOLTRRRPGQSKLSTPRDEKD 61  
Qy 114 TKILSGTHGVTGTTPTILVIVPNTDQSGDHREITANYRSHADATYDFKYGVRVAVQG 173  
Db 62 RVEIQSGTEFGKLTGTPIAMMIKNDQRPDHYSDMDKFPSPSHADFTYSEKYGKASSGG 121  
Qy 174 GRSSGRKTVGRAAGALPKILKLCGLEILSFVSKVHVVLPEDAVD-----YGSVTL 227  
Db 122 GRASARETIGRVAAGAIKFLAQNVEIVAFVTOIGEIKNRDSDPFBQHLNLTIR 181  
Qy 228 EQIES-NIVRCPDPEYAEKMDAIDRVRVSGDSVGGVITCVARNVPRGLSGSPVDFKLESE 286  
Db 182 EKVDWMPGIRCPDASVAGLWKEIEKYRGNKDSIGSVVTCVVRNLTGGLGPCFDKLEAM 241  
Qy 287 LAKAMLSIPASNGFPIGSGFAGTDLTGSEHNDEFYMDK-AGSVRTRTRSGVQGGISNV 345  
Db 242 LAHAMLSIPASKFPIGSGFQVSGVPSGKHNDPPYFEKETNRLRTKTNNSGVQGGISNG 301  
Qy 346 EIVHFKVAFKPTPSIGVKQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVLMQ 405  
Db 302 ENIYFSVPFKSVATISQOKTATYDGBEGILAAKGRHDPVTPRAIPIVEAMTALVLADA 361  
Qy 406 LMAHVAQ 412  
Db 362 LLIQKAR 368

RESULT 11  
US-09-103-331-40  
; Sequence 40, Application US/09103331  
; Patent No. 6699654  
; GENERAL INFORMATION:  
; APPLICANT: McLeod, Rima W.  
; APPLICANT: Roberts, Craig W.  
; APPLICANT: Roberts, Fiona  
; APPLICANT: Johnson, Jennifer J.  
; APPLICANT: Metz, Laurens J.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC  
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSES: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 23-JUN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Martin, Alice O.  
; REGISTRATION NUMBER: 35,601  
; REFERENCE/DOCKET NUMBER: 8336/9  
; TELEPHONE: 312-321-4200  
; TELEFAX: 312-321-4299  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 432 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-103-331-40

Query Match 42.3%; Score 930; DB 4; Length 432;  
Best Local Similarity 48.9%; Pred. No. 7.3e-94;  
Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;  
Qy 54 NTFGNVQVATYGESHGGVGVCGVISPPIPIPTLADLQVELDRRRPGQSRITSTRKETD 113  
Db 2 STFGHYFRVTTYGESHCKSVGCI VDGVPVPGMELTEDDIQPMTRRRPGQSAITTPRDEKD 61  
Qy 114 TKILSGTHGVTGTTPTILVIVPNTDQISGDH-REIANVY-RPSHADATYDFKYGVRVAVQ 171  
Db 62 RVIIFSQTERTGVTLTGTPIGMLVMNEDQPKDYGNKMTMDIYRPSHADWTYLEKYGVRKASS 121  
Qy 172 GGGSSGRKTVGRAAGALPKILKLCGLEILSFVSKVHVVL-----PED 218  
Db 122 GGGSSARETIGRVAAGAIKFLAQNVEIVAFVTOIGEIKNRDSDPFBQHLNLTIR 181  
Qy 219 ADVYGSVTLEQIESNI-VRCPDPEYAEKMDAIDRVRVSGDSVGGVITCVARNVPRGLGS 277  
Db 182 LKLVNSITRETVDVDFVRCFPAEANKMEDLITKFPDNHDSIGGTVTCVIRNVPSGLGE 241  
Qy 278 PVDFKLESELAAMLSIPASNGFPIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327  
Db 242 PAFDKLEAMLAHAMLSIPATKGFVSGFGCGCEVPSGIHNDPPFVSAENTEIPPSVAASGA 301  
Qy 328 VR-----TRNRSRGVQGGISNVEIVHFKVAFKPTPSIGVKQNTVSRERONVELL- 377  
Db 302 ARNGIPRPKUTTTNFGSGIQQGSGINAGPIYRFGFKPAATIGOEQTATYDGTSGVLA 361  
Qy 378 ARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFALNTALQBPVGS 428  
Db 362 AKGRHDSVVPRAVPIVEAMALVIMDAVLAHEARVTAKSLPLPKQTINS 412

RESULT 12  
US-09-631-594-49  
; Sequence 49, Application US/09631594  
; Patent No. 6737237  
; GENERAL INFORMATION:  
; APPLICANT: McLEOD, RIMA W.  
; APPLICANT: ROBERTS, CRAIG W.  
; APPLICANT: ROBERTS, FIONA  
; APPLICANT: JOHNSON, JENNIFER J.  
; APPLICANT: KRISITS, MICHAEL  
; APPLICANT: FERGUSON, DAVID  
; APPLICANT: LYONS, RUSSELL  
; APPLICANT: MUL, ERNEST  
; APPLICANT: MACK, DOUG  
; APPLICANT: SAMUEL, BENJAMIN  
; APPLICANT: GORNICKI, PIOTR  
; APPLICANT: ZUTHER, ELLEN  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES  
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS  
; FILE REFERENCE: 19338-90966  
; CURRENT APPLICATION NUMBER: US/09/631,594  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/US00/11478  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US97/12497  
; PRIOR FILING DATE: 1997-07-18  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 432  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
US-09-631-594-49

Query Match 42.3%; Score 930; DB 4; Length 432;  
Best Local Similarity 48.9%; Pred. No. 7.3e-94;  
Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;  
Qy 54 NTFGNVQVATYGESHGGVGVCGVISPPIPIPTLADLQVELDRRRPGQSRITSTRKETD 113

Db	2	STFGHYFRVTTYGESHCKSVG	ICVDGVP	GMBELTEDDIQ	QPMTRRRPGOSAITTPRDEK	61	
Qy	114	TKILSGTHEGVTGTPILVIVP	NTDQIGSDH-REI	ANVY-RP	SHADATYDFEKYGVRAVQ	171	
Db	62	RVIIQSGTFEGVTGTPICGL	VMNEDQPKDYGNK	TMDIYPR	PSHADWTYLEKYGVKASS	121	
Qy	172	GGGRSSGRKTVGRVNAAGAL	PKKILKGL	ELETLSFVSKVHQVL	-----PBD	218	
Db	122	GGGRSSARETIGRVAAGAIA	EKYLKPRYGE	IVAFVSVSGSHLFPPT	AEHPSPSPTNPEF	181	
Qy	219	ADVYGSVTLQIGTESNI-V	RCDDPEYAEKMI	DAIDRV	VRGDSVGVGITCVARNVPRGLCS	277	
Db	182	LKVNSITRETVDSEFLP	VRCPDAE	ANKMEDLI	TKFRDNHDSIGGTVCVIRNVPSGLGE	241	
Qy	278	PVPDKLESELAKAMLSI	PASNGFEH	IGSGFAGTDLTG	SEHNDEFY-----MDKAGS	327	
Db	242	PAFDKLEAMLAHAMLSI	PATKGF	FEVSGSGE	CVFPGSIHNDPFSVAENETI	PPSVAASGA	301
Qy	328	VR-----	TRTVRS	CGVQGGISN	VEIIVHFKVAFKTP	PSIGVKONTVSRERQNVELL-	377
Db	302	ARNGIPRPKLT	TKTNFSGGI	QGGISNG	APIYFRVCGFKPA	ATIGQSGTQATYDCTSEGVL	361
Qy	378	ARGHRDPCVAPRAVPV	VESMAALVLMQ	OLMAHVAQCEN	FALNTALOE	PVGS	428
Db	362	AKGHRDPSVPRAPVI	VEAMAA	VTMDRAVLA	HEARVTA	KSLPPKOTINS	412

RESULT 13  
US-09-489-039A-12546  
; Sequence 12546, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON ET. AL  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12546  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12546

```

      290 GGILGGISSGQOIVANIALKPTSSITVPFGHTINRFGEVEVMTKGRHDPVCVGIKRAVPAE 349
      396 SMAALVLMQDLMAHVAQ 412
      500 ANLAIVLMDHFMQRQA 560
      600 ANLAIVLMDHFMQRQA 660
      700 ANLAIVLMDHFMQRQA 720
      800 ANLAIVLMDHFMQRQA 780
      900 ANLAIVLMDHFMQRQA 840
      1000 ANLAIVLMDHFMQRQA 900
      1100 ANLAIVLMDHFMQRQA 960
      1200 ANLAIVLMDHFMQRQA 1020
      1300 ANLAIVLMDHFMQRQA 1080
      1400 ANLAIVLMDHFMQRQA 1140
      1500 ANLAIVLMDHFMQRQA 1200
      1600 ANLAIVLMDHFMQRQA 1260
      1700 ANLAIVLMDHFMQRQA 1320
      1800 ANLAIVLMDHFMQRQA 1380
      1900 ANLAIVLMDHFMQRQA 1440
      2000 ANLAIVLMDHFMQRQA 1500
      2100 ANLAIVLMDHFMQRQA 1560
      2200 ANLAIVLMDHFMQRQA 1620
      2300 ANLAIVLMDHFMQRQA 1680
      2400 ANLAIVLMDHFMQRQA 1740
      2500 ANLAIVLMDHFMQRQA 1800
      2600 ANLAIVLMDHFMQRQA 1860
      2700 ANLAIVLMDHFMQRQA 1920
      2800 ANLAIVLMDHFMQRQA 1980
      2900 ANLAIVLMDHFMQRQA 2040
      3000 ANLAIVLMDHFMQRQA 2100
      3100 ANLAIVLMDHFMQRQA 2160
      3200 ANLAIVLMDHFMQRQA 2220
      3300 ANLAIVLMDHFMQRQA 2280
      3400 ANLAIVLMDHFMQRQA 2340
      3500 ANLAIVLMDHFMQRQA 2400
      3600 ANLAIVLMDHFMQRQA 2460
      3700 ANLAIVLMDHFMQRQA 2520
      3800 ANLAIVLMDHFMQRQA 2580
      3900 ANLAIVLMDHFMQRQA 2640
      4000 ANLAIVLMDHFMQRQA 2700
      4100 ANLAIVLMDHFMQRQA 2760
      4200 ANLAIVLMDHFMQRQA 2820
      4300 ANLAIVLMDHFMQRQA 2880
      4400 ANLAIVLMDHFMQRQA 2940
      4500 ANLAIVLMDHFMQRQA 3000
      4600 ANLAIVLMDHFMQRQA 3060
      4700 ANLAIVLMDHFMQRQA 3120
      4800 ANLAIVLMDHFMQRQA 3180
      4900 ANLAIVLMDHFMQRQA 3240
      5000 ANLAIVLMDHFMQRQA 3300
      5100 ANLAIVLMDHFMQRQA 3360
      5200 ANLAIVLMDHFMQRQA 3420
      5300 ANLAIVLMDHFMQRQA 3480
      5400 ANLAIVLMDHFMQRQA 3540
      5500 ANLAIVLMDHFMQRQA 3600
      5600 ANLAIVLMDHFMQRQA 3660
      5700 ANLAIVLMDHFMQRQA 3720
      5800 ANLAIVLMDHFMQRQA 3780
      5900 ANLAIVLMDHFMQRQA 3840
      6000 ANLAIVLMDHFMQRQA 3900
      6100 ANLAIVLMDHFMQRQA 3960
      6200 ANLAIVLMDHFMQRQA 4020
      6300 ANLAIVLMDHFMQRQA 4080
      6400 ANLAIVLMDHFMQRQA 4140
      6500 ANLAIVLMDHFMQRQA 4200
      6600 ANLAIVLMDHFMQRQA 4260
      6700 ANLAIVLMDHFMQRQA 4320
      6800 ANLAIVLMDHFMQRQA 4380
      6900 ANLAIVLMDHFMQRQA 4440
      7000 ANLAIVLMDHFMQRQA 4500
      7100 ANLAIVLMDHFMQRQA 4560
      7200 ANLAIVLMDHFMQRQA 4620
      7300 ANLAIVLMDHFMQRQA 4680
      7400 ANLAIVLMDHFMQRQA 4740
      7500 ANLAIVLMDHFMQRQA 4800
      7600 ANLAIVLMDHFMQRQA 4860
      7700 ANLAIVLMDHFMQRQA 4920
      7800 ANLAIVLMDHFMQRQA 4980
      7900 ANLAIVLMDHFMQRQA 5040
      8000 ANLAIVLMDHFMQRQA 5100
      8100 ANLAIVLMDHFMQRQA 5160
      8200 ANLAIVLMDHFMQRQA 5220
      8300 ANLAIVLMDHFMQRQA 5280
      8400 ANLAIVLMDHFMQRQA 5340
      8500 ANLAIVLMDHFMQRQA 5400
      8600 ANLAIVLMDHFMQRQA 5460
      8700 ANLAIVLMDHFMQRQA 5520
      8800 ANLAIVLMDHFMQRQA 5580
      8900 ANLAIVLMDHFMQRQA 5640
      9000 ANLAIVLMDHFMQRQA 5700
      9100 ANLAIVLMDHFMQRQA 5760
      9200 ANLAIVLMDHFMQRQA 5820
      9300 ANLAIVLMDHFMQRQA 5880
      9400 ANLAIVLMDHFMQRQA 5940
      9500 ANLAIVLMDHFMQRQA 6000
      9600 ANLAIVLMDHFMQRQA 6060
      9700 ANLAIVLMDHFMQRQA 6120
      9800 ANLAIVLMDHFMQRQA 6180
      9900 ANLAIVLMDHFMQRQA 6240
      10000 ANLAIVLMDHFMQRQA 6300
      10100 ANLAIVLMDHFMQRQA 6360
      10200 ANLAIVLMDHFMQRQA 6420
      10300 ANLAIVLMDHFMQRQA 6480
      10400 ANLAIVLMDHFMQRQA 6540
      10500 ANLAIVLMDHFMQRQA 6600
      10600 ANLAIVLMDHFMQRQA 6660
      10700 ANLAIVLMDHFMQRQA 6720
      10800 ANLAIVLMDHFMQRQA 6780
      10900 ANLAIVLMDHFMQRQA 6840
      11000 ANLAIVLMDHFMQRQA 6900
      11100 ANLAIVLMDHFMQRQA 6960
      11200 ANLAIVLMDHFMQRQA 7020
      11300 ANLAIVLMDHFMQRQA 7080
      11400 ANLAIVLMDHFMQRQA 7140
      11500 ANLAIVLMDHFMQRQA 7200
      11600 ANLAIVLMDHFMQRQA 7260
      11700 ANLAIVLMDHFMQRQA 7320
      11800 ANLAIVLMDHFMQRQA 7380
      11900 ANLAIVLMDHFMQRQA 7440
      12000 ANLAIVLMDHFMQRQA 7500
      12100 ANLAIVLMDHFMQRQA 7560
      12200 ANLAIVLMDHFMQRQA 7620
      12300 ANLAIVLMDHFMQRQA 7680
      12400 ANLAIVLMDHFMQRQA 7740
      12500 ANLAIVLMDHFMQRQA 7800
      12600 ANLAIVLMDHFMQRQA 7860
      12700 ANLAIVLMDHFMQRQA 7920
      12800 ANLAIVLMDHFMQRQA 7980
      12900 ANLAIVLMDHFMQRQA 8040
      13000 ANLAIVLMDHFMQRQA 8100
      13100 ANLAIVLMDHFMQRQA 8160
      13200 ANLAIVLMDHFMQRQA 8220
      13300 ANLAIVLMDHFMQRQA 8280
      13400 ANLAIVLMDHFMQRQA 8340
      13500 ANLAIVLMDHFMQRQA 8400
      13600 ANLAIVLMDHFMQRQA 8460
      13700 ANLAIVLMDHFMQRQA 8520
      13800 ANLAIVLMDHFMQRQA 8580
      13900 ANLAIVLMDHFMQRQA 8640
      14000 ANLAIVLMDHFMQRQA 8700
      14100 ANLAIVLMDHFMQRQA 8760
      14200 ANLAIVLMDHFMQRQA 8820
      14300 ANLAIVLMDHFMQRQA 8880
      14400 ANLAIVLMDHFMQRQA 8940
      14500 ANLAIVLMDHFMQRQA 9000
      14600 ANLAIVLMDHFMQRQA 9060
      14700 ANLAIVLMDHFMQRQA 9120
      14800 ANLAIVLMDHFMQRQA 9180
      14900 ANLAIVLMDHFMQRQA 9240
      15000 ANLAIVLMDHFMQRQA 9300
      15100 ANLAIVLMDHFMQRQA 9360
      15200 ANLAIVLMDHFMQRQA 9420
      15300 ANLAIVLMDHFMQRQA 9480
      15400 ANLAIVLMDHFMQRQA 9540
      15500 ANLAIVLMDHFMQRQA 9600
      15600 ANLAIVLMDHFMQRQA 9660
      15700 ANLAIVLMDHFMQRQA 9720
      15800 ANLAIVLMDHFMQRQA 9780
      15900 ANLAIVLMDHFMQRQA 9840
      16000 ANLAIVLMDHFMQRQA 9900
      16100 ANLAIVLMDHFMQRQA 9960
      16200 ANLAIVLMDHFMQRQA 10020
      16300 ANLAIVLMDHFMQRQA 10080
      16400 ANLAIVLMDHFMQRQA 10140
      16500 ANLAIVLMDHFMQRQA 10200
      16600 ANLAIVLMDHFMQRQA 10260
      16700 ANLAIVLMDHFMQRQA 10320
      16800 ANLAIVLMDHFMQRQA 10380
      16900 ANLAIVLMDHFMQRQA 10440
      17000 ANLAIVLMDHFMQRQA 10500
      17100 ANLAIVLMDHFMQRQA 10560
      17200 ANLAIVLMDHFMQRQA 10620
      17300 ANLAIVLMDHFMQRQA 10680
      17400 ANLAIVLMDHFMQRQA 10740
      17500 ANLAIVLMDHFMQRQA 10800
      17600 ANLAIVLMDHFMQRQA 10860
      17700 ANLAIVLMDHFMQRQA 10920
      17800 ANLAIVLMDHFMQRQA 10980
      17900 ANLAIVLMDHFMQRQA 11040
      18000 ANLAIVLMDHFMQRQA 11100
      18100 ANLAIVLMDHFMQRQA 11160
      18200 ANLAIVLMDHFMQRQA 11220
      18300 ANLAIVLMDHFMQRQA 11280
      18400 ANLAIVLMDHFMQRQA 11340
      18500 ANLAIVLMDHFMQRQA 11400
      18600 ANLAIVLMDHFMQRQA 11460
      18700 ANLAIVLMDHFMQRQA 11520
      18800 ANLAIVLMDHFMQRQA 11580
      18900 ANLAIVLMDHFMQRQA 11640
      19000 ANLAIVLMDHFMQRQA 11700
      19100 ANLAIVLMDHFM
```

```

RESULT 15
US-09-103-331-41
; Sequence 41, Application US/09103331
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Mets, Laurens J.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE
; APICOMPLEXAN

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:28:51 ; Search time 161 Seconds  
(without alignments)  
1046.990 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQAHAHRLAPR.....VAQCEMFALNTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	429	15	US-10-677-179-8
2	2107	95.8	497	16	US-10-425-115-323000
3	2016.5	91.7	408	15	US-10-425-114-69504
4	2011.5	91.4	402	16	US-10-425-115-322998
5	1800.5	81.8	440	16	US-10-425-115-356153
6	1797.5	81.7	440	15	US-10-677-179-2
7	1790.5	81.4	452	15	US-10-425-114-57857
8	1724	78.4	426	16	US-10-437-963-201690
9	1541	70.0	431	14	US-10-268-611-39
10	1541	70.0	431	18	US-10-824-194-48
11	1527	69.4	440	14	US-10-268-611-63

12	1512	68.7	435	15	US-10-424-599-144653
13	1512	68.7	438	15	US-10-425-114-52276
14	1454	66.1	436	16	US-10-636-616-28
15	1257	57.1	307	15	US-10-677-179-12
16	1191.5	54.2	362	14	US-10-268-611-38
17	1191.5	54.2	362	14	US-10-268-611-64
18	1191.5	54.2	362	15	US-10-369-493-2677
19	1191.5	54.2	362	18	US-10-824-194-47
20	1186.5	53.9	362	15	US-10-369-493-18868
21	1182.5	53.8	362	15	US-10-369-493-20210
22	1066	48.5	257	15	US-10-677-179-6
23	1035	47.0	360	15	US-10-369-493-21053
24	981	44.6	356	15	US-10-369-493-10419
25	938	42.6	376	14	US-10-268-611-42
26	938	42.6	376	15	US-10-369-493-21922
27	938	42.6	376	18	US-10-824-194-51
28	934	42.5	378	17	US-10-741-849-7185
29	930	42.3	432	14	US-10-268-611-40
30	930	42.3	432	18	US-10-824-194-49
31	927.5	42.2	285	15	US-10-424-599-144654
32	903	41.0	240	16	US-10-425-115-356150
33	888.5	40.4	358	15	US-10-282-122A-48831
34	886.5	40.3	473	15	US-10-369-493-3972
35	862	39.2	361	15	US-10-282-122A-78247
36	860	39.1	361	9	US-09-815-242-10232
37	860	39.1	361	15	US-10-369-493-23483
38	860	39.1	361	15	US-10-282-122A-56623
39	858	39.0	352	15	US-10-282-122A-61027
40	857	39.0	361	9	US-09-815-242-11708
41	856	38.9	361	14	US-10-268-611-65
42	855	38.9	361	15	US-10-282-122A-59797
43	850	38.6	361	15	US-10-282-122A-69065
44	845.5	38.5	357	15	US-10-282-122A-66978
45	843	38.3	360	15	US-10-369-493-21155

#### ALIGNMENTS

RESULT 1  
US-10-677-179-8  
; Sequence 8, Application US/10677179  
; Publication No. US20040082050A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/10/677,179  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US/09/743,207  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-677-179-8

Query Match 100.0%; Score 2200; DB 15; Length 429;  
Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTTVPKQQAHAHRLAPRAIGALLEFAFASLSRLFAVHRCRTARLEVKASGNTFGNYF 60  
Db 1 MTTVPKQQAHAHRLAPRAIGALLEFAFASLSRLFAVHRCRTARLEVKASGNTFGNYF 60

```
QY 61 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 120
Db 61 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 120
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 180
Db 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 240
Db 181 TVGRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 240
QY 241 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 241 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQVGGVSNVEIVHFVKVAFKPTPSI 360
Db 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQVGGVSNVEIVHFVKVAFKPTPSI 360
QY 361 GVKQNTVSRERONVELLAGRHDPCVAPRAVPVVSMAALVLMQJMAHVAQCENFALNT 420
Db 361 GVKQNTVSRERONVELLAGRHDPCVAPRAVPVVSMAALVLMQJMAHVAQCENFALNT 420
QY 421 ALQEPVGSF 429
Db 421 ALQEPVGSF 429
```

## RESULT 2

```
US-10-425-115-323000
; Sequence 323000, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 323000
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(497)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57640C.1.pep
US-10-425-115-323000
```

Query Match 95.8%; Score 2107; DB 16; Length 497;

Best Local Similarity 84.9%; Pred. No. 2.5e-200;

Matches 422; Conservative 2; Mismatches 5; Indels 68; Gaps 2;

```
QY 1 MTTVPKQQAHSRARLAPRAIGALLEFAPASSSLRFAVHRCRTARLEVKASGNTFGNYF 60
Db 1 MTTVPKQQAHSRARLAPRAIGALLEFAPASSSLRFAVHRCRTARLEVKASGNTFGNYF 60
QY 61 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 120
Db 61 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 120
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 180
Db 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 240
```

```
Db 181 TVGRVAAGALAKLILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 240
QY 241 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 241 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQVGGVSNVEIVHFVKVAFKPTPSI 337
Db 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQVGGVSNVEIVHFVKVAFKPTPSI 360
QY 338 -----VOGGISNVEIVHFVKVAFKPTPSIGVKQNTVSRERONVELLAGRHDPCVAPR 389
Db 361 NGIGLFXAVQGGISNGSIVHFVKVAFKPTPSIGVKQNTVSRERONVELLAGRHDPCVAPR 420
QY 390 -----AVPVESMAALVLMQJMAHVAQ 412
Db 421 XHVRVMTLLAASATILFNFPPATATATGIRGLLLTWLXAVPVESMAALVLMQJMAHVAQ 480
QY 413 CEMFALNTALQEPVGSF 429
Db 481 CEMFALNTALQEPVGSF 497

RESULT 3
US-10-425-114-69504
; Sequence 69504, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69504
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLE73045G02_FLI1.pep
US-10-425-114-69504
```

Query Match 91.7%; Score 2016.5; DB 15; Length 408;

Best Local Similarity 92.8%; Pred. No. 1.9e-191;

Matches 398; Conservative 1; Mismatches 3; Indels 27; Gaps 1;

```
QY 1 MTTVPKQQAHSRARLAPRAIGALLEFAPASSSLRFAVHRCRTARLEVKASGNTFGNYF 60
Db 7 MTTVPKQQAHSRARLAPRAIGALLEFAPASSSLRFAVHRCRTARLEVKASGNTFGNYF 66
QY 61 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 120
Db 67 QVATYGESHGGVGVISGCPPIPLTEADLQVELDRRRPGQSRTSTRKETDTCILSG 126
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 180
Db 127 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAVQGGSSGRK 186
QY 181 TVGRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 240
Db 187 TVGRVAAGALAKLILKLCGLLEILSFVSKVQHVLPEDAVDYGSTLQIESNIVRCDDP 219
QY 241 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 220 EYAEKMIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 279
```

Qy	301	ETGSGFAGTDLTGSEHNDEFYMDKAGSVRTNRSGGVGGISNVELVHFKVAFKPTPSI	360
Db	280	ETGSGFAGTDLTGSEHNDEFYMDKAGSVRTNRSGGVGGISNGEVLHFKVAFKPTPSI	339
Qy	361	GVKQNTVSRERQNVELLARGRHDPVAPRAVPVVSMAALVLMQDMAHVAQCEMPALNT	420
Db	340	GVKQNTVSRERQNVELLARGRHDPVAPRAVPVVSMAALVLMQDMAHVAQCEMPALNT	399
Qy	421	ALQEPVGSF	429
Db	400	ALQEPVGSF	408

```

RESULT 4
US-10-425-115-322998
; Sequence 322998, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 322998
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_57639C.1.pgp
US-10-425-115-322998

```

```

RESULT 5
US-10-425-115-356153
; Sequence 356153, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 356153
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_87981C.1.rep
US-10-425-115-356153

```

```

RESULT 6
US-10-677-179-2
; Sequence 2, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Canoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207

```

```
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Zea mays
US-10-677-179-2

Query Match      81.7%; Score 1797.5; DB 15; Length 440;
Best Local Similarity 81.9%; Pred. No. 1.3e-169;
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;

Qy  4 VPKPOQVAHSRARLAPRAIGALLEPAPASSLSRFV-HRCRTARLEVKASGNTFGNYFOV 62
Db  5 VSQPPVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRAASLEVKASGNVFGNYFOV 62
Qy  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Db  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Qy  123 EGVTTGTPILVIVPNTDOI GSDHREIANVVRPASHADATYDFKYGVRVAVQGGSSGRKTV 182
Db  123 DGMTTGTPIHVFVNTDQGGDYSEMSKAYRPSHADATYDFKYGVRVAVQGGSSARETI 182
Qy  183 GRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSVTLQEIENIVRCPDPEY 242
Db  183 GRVAAGALAKKILKLSGVLEILAFVSKVQHVLPEDAVDYETVTLHEIESNIVRCPDPEY 242
Qy  243 AEKMDAIDRVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEI 302
Db  243 AEKMTAAIDTVRVGRDSIGGVVTCIARNVPRGLSPVFDKLESELAKAMLSIPASKGFEI 302
Qy  303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSQVGGISNVEIVHFKVAFKPTPSIGV 362
Db  303 GSGFAGTDTFTGSEHNDEFYMDAGNVRTRTRNRSQVGGISNGEIIYFKVAFKPTATIGK 362
Qy  363 KONTVSRRQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTAL 422
Db  363 KONTVSREHEDVELLARGHDPVVPRAVPVVSMAALVMDQLMAHIAQCEMPPLNAL 422
Qy  423 QEPVGS 428
Db  423 QEPIGS 428

RESULT 7
US-10-425-114-57857
; Sequence 57857, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57857
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73278F07_FLI.pep

Query Match      78.4%; Score 1724; DB 16; Length 426;
Best Local Similarity 80.6%; Pred. No. 2.6e-162;
Matches 337; Conservative 31; Mismatches 32; Indels 18; Gaps 3;

Qy  15 ARLAPRAIG----ALLEFAPASSLSRFVHRCRTARLEVKASGNTFGNYFOVATYGESHG 70
Db  11 ARVLPRGGGGGFRAPPESAPA--SLRFSVGRRAARLEVKASANVFGNYFOVATYGESHG 68
```

```
US-10-425-114-57857

Query Match      81.4%; Score 1790.5; DB 15; Length 452;
Best Local Similarity 81.9%; Pred. No. 6.9e-169;
Matches 349; Conservative 31; Mismatches 43; Indels 3; Gaps 2;

Qy  4 VPKPOQVAHSRARLAPRAIGALLEPAPASSLSRFV-HRCRTARLEVKASGNTFGNYFOV 62
Db  17 VSQPPVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRAASLEVKASGNVFGNYFOV 74
Qy  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Db  75 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITPRKETDTCKILSGTH 134
Qy  123 EGVTTGTPILVIVPNTDOI GSDHREIANVVRPASHADATYDFKYGVRVAVQGGSSGRKTV 182
Db  135 DGMTTGTPIHVFVNTDQGGDYSEMSKAYRPSHADATYDFKYGVRVAVQGGSSARETI 194
Qy  183 GRVAAGALPKKILKLCGLLEILSFVSKVQHVLPEDAVDYGSVTLQEIENIVRCPDPEY 242
Db  195 GRVAAGALAKKILKLSGVLEILAFVSKVQHVLPEDAVDYETVTLHEIESNIVRCPDPEY 254
Qy  243 AEKMDAIDRVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEI 302
Db  255 AEKMTAAIDTVRVGRDSIGGVVTCIARNVPRGLSPVFDKLESELAKAMLSIPASKGFEI 314
Qy  303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSQVGGISNVEIVHFKVAFKPTPSIGV 362
Db  315 GSGFAGTDTFTGSEHNDEFYMDAGNVRTRTRNRSQVGGISNGEIIYFKVAFKPTATIGK 374
Qy  363 KONTVSRRQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTAL 422
Db  375 KONTVSREHEDVELLARGHDPVVPRAVPVVSMAALVMDQLMAHIAQCEMPPLNAL 434
Qy  423 QEPVGS 428
Db  435 QEPIGS 440

RESULT 8
US-10-437-963-201690
; Sequence 201690, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 201690
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9703C.1.pep
US-10-437-963-201690

Query Match      78.4%; Score 1724; DB 16; Length 426;
Best Local Similarity 80.6%; Pred. No. 2.6e-162;
Matches 337; Conservative 31; Mismatches 32; Indels 18; Gaps 3;

Qy  15 ARLAPRAIG----ALLEFAPASSLSRFVHRCRTARLEVKASGNTFGNYFOVATYGESHG 70
Db  11 ARVLPRGGGGGFRAPPESAPA--SLRFSVGRRAARLEVKASANVFGNYFOVATYGESHG 68
```



QY 71 GGVGCVISGPPRIPLTEADLQVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTP 130  
DB 69 GGVGCVISGPPRIPLTEADLQVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTP 128  
QY 131 ILVIVNTDQIGSDHREIANVYRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGAL 190  
DB 129 IHVFEVNTDQRGDYSMAKAYRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGAL 188  
QY 191 PKILKLCGLLEILSPVSKVHVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAI 250  
DB 189 AKILKLCGLLEILSPVSKVHVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAI 248  
QY 251 DRVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTD 310  
DB 249 DKVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTD 308  
QY 311 LTGSEHNDFFYMDKAGSVTRTRNSGVQGGISNIVHFKVAFKPTPSIGVKQNTVSRE 370  
DB 309 YTGSEHNDFFYMDAGN-----GGISNGSIIVFKVAFKPTATIGKKQHTVSRE 356  
QY 371 RONVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAOCMEPALNTALQEPVGS 428  
DB 357 HEDVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAOCMEPALNTALQEPVGS 414

## RESULT 9

US-10-268-611-39  
; Sequence 39, Application US/10268611  
; Publication No. US20030186352A1  
; GENERAL INFORMATION:  
; APPLICANT: MCLEOD, RIMA W.  
; APPLICANT: ROBERTS, CRAIG W.  
; APPLICANT: ROBERTS, FIONA  
; APPLICANT: JOHNSON, JENNIFER J.  
; APPLICANT: KIRISITS, MICHAEL  
; APPLICANT: FERGUSON, DAVID  
; APPLICANT: LYONS, RUSSELL  
; APPLICANT: MUI, ERNEST  
; APPLICANT: MACK, DOUG  
; APPLICANT: SAMUEL, BENJAMIN  
; APPLICANT: GORNICKI, PIOTR  
; APPLICANT: ZUTHER, ELLEN  
; TITLE OF INVENTION: APICOMPLEXAN CHORISMATE SYNTHASE SEQUENCES AND AN  
; FILE REFERENCE: 19338/93888  
; CURRENT APPLICATION NUMBER: US/10/268,611  
; PRIOR FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: 09/103,331  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/329,269  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/346,036  
; PRIOR FILING DATE: 2001-11-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 39  
; TYPE: PRT  
; LENGTH: 431  
; ORGANISM: Solanum lycopersicum  
US-10-268-611-39

Query Match 70.0%; Score 1541; DB 14; Length 431;  
Best Local Similarity 72.7%; Pred. No. 4.4e-144;  
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 32 SSSLRFVAVHRCRTARLEVKASGNTFGNYFQVATYGESHGGVGVISGCPPIPLTEADL 91  
DB 29 SSNLFPFTHRSQPKLEIQAGNFGNYFRVTTFGESHGGVGVICIDGCPPLPLSESDM 88  
QY 92 QVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTPILVIVNTDQIGSDHREIANV 151  
DB 89 QVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTPILVIVNTDQIGSDHREIANV 148

QY 152 YRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSPVSKVH 211  
DB 149 YRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSPVSKVH 208  
QY 212 QVVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAIDRVRVRGDSVGGVITCVARNV 271  
DB 209 NVVLPEDLVDNQIVTLQIESNIVRCPPPEYAEKIMDAIDRVRVRGDSVGGVITCVARNV 268  
QY 272 PRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDFFYMDKAGSVTR 331  
DB 269 PRGLGTPTVPFDKLESELAKAMLSIPATKGFEGSGFAGTGMTGSEHNDFFYMDHDIQIRTK 328  
QY 332 TNRSVGQGGISNIVHFKVAFKPTPSIGVKQNTVSREQNVVELLARGHDPVAPRAV 391  
DB 329 TNRSVGQGGISNIVHFKVAFKPTPSIGVKQNTVSREQNVVELLARGHDPVAPRAV 388  
QY 392 PVVESMAALVMDQLMAHVAOCMEPALNTALQEPV 426  
DB 389 PVVEAMVALVLDQLMTQVACMLFPVNLTLQEP 423

## RESULT 10

US-10-824-194-48  
; Sequence 48, Application US/10824194  
; Publication No. US20050142113A1  
; GENERAL INFORMATION:  
; APPLICANT: MCLEOD, RIMA W.  
; APPLICANT: ROBERTS, CRAIG W.  
; APPLICANT: ROBERTS, FIONA  
; APPLICANT: JOHNSON, JENNIFER J.  
; APPLICANT: KIRISITS, MICHAEL  
; APPLICANT: FERGUSON, DAVID  
; APPLICANT: LYONS, RUSSELL  
; APPLICANT: MUI, ERNEST  
; APPLICANT: MACK, DOUG  
; APPLICANT: SAMUEL, BENJAMIN  
; APPLICANT: GORNICKI, PIOTR  
; APPLICANT: ZUTHER, ELLEN  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES  
; FILE REFERENCE: 19338-90966  
; CURRENT APPLICATION NUMBER: US/10/824,194  
; PRIOR FILING DATE: 2004-04-14  
; PRIOR APPLICATION NUMBER: US/09/631,594  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/US00/11478  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US97/12497  
; PRIOR FILING DATE: 1997-07-18  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Solanum lycopersicum  
US-10-824-194-48

Query Match 70.0%; Score 1541; DB 18; Length 431;  
Best Local Similarity 72.7%; Pred. No. 4.4e-144;  
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 32 SSSLRFVAVHRCRTARLEVKASGNTFGNYFQVATYGESHGGVGVISGCPPIPLTEADL 91  
DB 29 SSNLFPFTHRSQPKLEIQAGNFGNYFRVTTFGESHGGVGVICIDGCPPLPLSESDM 88  
QY 92 QVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTPILVIVNTDQIGSDHREIANV 151  
DB 89 QVELDRRPGQSRITSTRKETDTCKILSGTHEGVTGTPILVIVNTDQIGSDHREIANV 148  
QY 152 YRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSPVSKVH 211  
DB 149 YRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSPVSKVH 208

Qy 212 QVLPEDADVXSVTLQIESNIVRCPDPEYAEKMDAIDRVVRVGRDVGSGVITCVARNV 271  
Db 209 NVVLPEDLVNDQIVTLQIESNIVRCPNPEYAEKMGDAIDRVVRVGRDVGSGVITCVARNV 268  
Qy 272 PRGLGSPVFDKLESELAKAMLSIPASNGEIGSGFAGTDLTGSHEHNDFFYMDKAGSVRTR 331  
Db 269 PRGLGTPVFDKLESELAKACHSLPATKGFEGSGFAGTFMTGSEHNDFFWDEHDQIRTK 328  
Qy 332 TNRSQGVGGISNIVBIHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPCCVAPRAV 391  
Db 329 TNRSQGIQGGISNGEIIINRVAFKPTSTIARKQHTVSRDKHETELIARGRHDPCCVAPRAV 388  
Qy 392 PVVESMAALVMDQLMAHVACCEMPALNTALQEPV 426  
Db 389 PMVEAMVALVLDQLMTQYAOQMLPFVNLTLQEPV 423

## RESULT 11

US-10-268-611-63

; Sequence 63, Application US/10268611

; Publication No. US20030186352A1

; GENERAL INFORMATION:

; APPLICANT: MCLEOD, RIMA W.

; APPLICANT: ROBERTS, CRAIG W.

; APPLICANT: ROBERTS, FIONA

; APPLICANT: JOHNSON, JENNIFER J.

; APPLICANT: KIRISITS, MICHAEL

; APPLICANT: FERGUSON, DAVID

; APPLICANT: LYONS, RUSSELL

; APPLICANT: MUI, ERNEST

; APPLICANT: MACK, DOUG

; APPLICANT: SAMUEL, BENJAMIN

; APPLICANT: GORNICKI, PIOTR

; APPLICANT: ZUTHER, ELEN

; TITLE OF INVENTION: APICOMPLEXAN CHORISMATE SYNTHASE SEQUENCES AND AN

; FILE REFERENCE: 19338/93888

; CURRENT APPLICATION NUMBER: US/10/268,611

; CURRENT FILING DATE: 2002-10-10

; PRIOR APPLICATION NUMBER: 09/103,331

; PRIOR FILING DATE: 1998-06-23

; PRIOR APPLICATION NUMBER: 60/329,269

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/346,036

; PRIOR FILING DATE: 2001-11-08

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 63

; LENGTH: 440

; TYPE: PRT

; ORGANISM: Solanum lycopersium

US-10-268-611-63

## Query Match

Best Local Similarity 69.4%; Score 1527; DB 14; Length 440;

Matches 288; Conservative 64; Mismatches 73; Indels 2; Gaps 1;

Qy 4 VPKPQQVAHSRAR--LAPRAIGALLEFAPASSSLRFVHRCRTARLEVKASGNTFGNYFQ 61  
Db 5 VPTKQFVGASSSSDGGSLVLSQLPSKFSNFFLPSRPSQLKRLQIQAAGSTFGNYFR 64  
Qy 62 VATYGESHGGVGVISGCPPIPLTEADLOVELDRRRPGQSRTSTRKETDTCILSGT 121  
Db 65 VTTGESHGGVGVICIDGCPPLPSLSDMQVELDRRRPGQSRTSTRKETDTCISGT 124  
Qy 122 HEGVTTGTPILVINTQIGSDHREIANVYRP SHADATYDPKYGVRVAVQGGSSGRKT 181  
Db 125 ADGLTTSFPIKEVNTQORNDYSEMSLAYRPSHADATYDPKYGVRVAVQGGSSGRKT 184  
Qy 182 GVRVAAGALPKKILKCGLEILSFVSKVHVLPEDAVDYGSVTLQIESNIVRCPDPE 241  
Db 185 IGRVAAGA VAKKILKLYSGAEVLAVVSVQHVLPEDLIDHQNVTLEQIESNIVRCPDPE 244

Qy 242 YAEKMDAIDRVVRVGRDVGSGVITCVARNVPRGIGSPVFDKLESELAKAMLSIPASNGPE 301  
Db 245 YAEKMDAIDRVVRVGRDVGSGVITCVARNVPRGIGTPVFDKLESELAKACHSLPATKAGFE 304  
Qy 302 IGSFAGTDLTGSHEHNDFFYMDKAGSVRTRTNRSQGVGGISNVEIVHFVKVAFKPTPSIG 361  
Db 305 FGSFAGTFTGTGSEHNDFFYMDHGRIRTRTNRSQGIQGGISNGEIVNMRIGPKPTSTIS 364  
Qy 362 VKQNTVSRERQNVLLARGRHDPCCVAPRAVPVSVESMAALVMDQLMAHVACCEMPALNTA 421  
Db 365 RKQQTVTRDKHETELIARGRHDPCCVAPRAVPVMEAMVALVLDQLMAQYSCQMMFPINPE 424  
Qy 422 LQEPVGS 428  
Db 425 LQEPVGS 431

## RESULT 12

US-10-424-599-144653

; Sequence 144653, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 144653

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101634C.1.pap

US-10-424-599-144653

## Query Match

Best Local Similarity 68.7%; Score 1512; DB 15; Length 435;

Matches 283; Conservative 62; Mismatches 63; Indels 8; Gaps 1;

Qy 19 PRAIGALLEFAPASS-----LRFVHRCRTARLEVKASGNTFGNYFQVATYGESHG 70  
Db 9 PFSADALSASFASLNSDLGSLSPAYLRSLRPLRPKRLRIQAAGSTYGNHFRVTTYGESHG 68  
Qy 71 GGVCVVISGCPPIPLTEADLOVELDRRRPGQSRTSTRKETDTCILSGTHEGVTTGTP 130  
Db 69 GGVCVIDGCPPLPLSEADMQVLDLDRRRPGQSRTTTPRKETDTCIFSGVSEGITTTGTP 128  
Qy 131 ILVTVPNTDQIGSDHREIANVYRP SHADATYDPKYGVRVAVQGGSSGRKTGVRVAAGAL 190  
Db 129 IHVSVPNTDQIGSDHREIANVYRP SHADATYDPKYGVRVAVQGGSSGRKTGVRVAAGAL 188  
Qy 191 PKKILKLCGLEILSFVSKVHVLPEDAVDYGSVTLQIESNIVRCPDPEYAEKMDAI 250  
Db 189 AKKILKEFSGTEILAVVSVQHKIVLPEDLIDHDTLTDQIESNIVRCPDPEYAEKMSAI 248  
Qy 251 DRVVRVGRDVGSGVITCVARNVPRGIGSPVFDKLESELAKAMLSIPASNGPEIGSGFAGTD 310  
Db 249 DAVVRVGRDVGSGVITCVARNVPRGIGSPVFDKLESELAKAMLSIPATKGFQFGSGFAGTF 308  
Qy 311 LTGSEHNDFFYMDKAGSVRTRTNRSQGVGGISNVEIVHFVKVAFKPTPSIGVKQNTVSR 370  
Db 309 LTGSEHNDFFYMDHGRIRTRTNRSQGIQGGISNGEIVNMRIGPKPTSTIS 368  
Qy 371 QNVVLLARGRHDPCCVAPRAVPVSVESMAALVMDQLMAHVACCEMPALNTALQEPV 426  
Db 369 KKETEFTIARGRHDPCCVAPRAVPVMEAMVALVLDQLMAQYSCQMMFPINPE 424

RESULT 13  
US-10-425-114-52276  
; Sequence 52276, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 28  
; SEQ ID NO 52276  
; LENGTH: 438  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700829731\_FLI.pap  
US-10-425-114-52276

Query Match 68.7%; Score 1512; DB 15; Length 438;  
Best Local Similarity 68.0%; Pred. No. 3.5e-141;  
Matches 283; Conservative 62; Mismatches 63; Indels 8; Gaps 1;  
QY 19 PRAIGALLEFAPASS-----LRFVHRCRTARLEVKASGNTFGNYFQVATYGESHG 70  
DB 12 PFSADALSASFSLNSDLGSLPPAYLRLSLRPLPKLHLQAAGSTVGNHFRVTYGESHG 71  
QY 71 GGVGVISCPRIPLTEADLQVELDRRPPGQSRTSTKTDTCILSGTHEGVTGTPILVIVPNTDQIGSDHREIAN 130  
DB 72 GGVGVICDPCPLPLEADMQVDLDRRPPGQSRTSTKTDTCILSGTHEGVTGTP 131  
QY 131 ILVIVPNTDQIGSDHREIANVPPSHADATYFKYGVRAVOGGSSGRKTVGRVAAGAL 190  
DB 132 IHVSVENTDQRGHDYSEMAVAPRPSHADATYDMKYGVRSVOGGSSARETIGRVASGAV 191  
QY 191 PKKILKLCGLLEILSVSKVHVQVLPEDAVDYGSTVLEQIESNIVRCPPPEYAEKMDAI 250  
DB 192 AKKILKEPSGTEILAVYSQVHKVLEPDLIDHTLTLDQIESNIVRCPPPEYAEKMDAI 251  
QY 251 DRVRGDSVGVVITCVARNVRGLGSPVFDKLESELAKAMLSIPASNGFBIIGSGFAGTD 310  
DB 252 DAVVRGDSVGVVITCVARNVRGLGSPVFDKLESELAKAMLSIPATKGFQFGSGFAGTF 311  
QY 311 LNDSEHNDEFYMDKAGSVTRTRNSGGVGGISNVEIHFVKAFKPTPSIGVKONTVSRRE 370  
DB 312 LTGSEHNDEFYIDEGHNTTRTRNSGGIQQGINSGEIINMRVAFKPTSTIGKKQKTVTRD 371  
QY 371 RQVVELLARGHDPVAPRAVPVBSMAALVMDQLMAHVAOCMEFALNTALQEPV 426  
DB 372 KKSETFIARGHDPVAPRAVPVBSMAALVMDQLMAHVAOCMEFALNTALQEPV 427

RESULT 14  
US-10-696-616-28  
; Sequence 28, Application US/10696616  
; Publication No. US20040191850A1  
; GENERAL INFORMATION:  
; APPLICANT: Levin, Joshua  
; APPLICANT: Tossberg, John  
; APPLICANT: Zhou, Qing  
; APPLICANT: McElver, John  
; APPLICANT: Aux, George  
; APPLICANT: Budziszewski, Greg  
; APPLICANT: Thomas, Carla  
; APPLICANT: Patton, David  
; APPLICANT: Frye, Catherine

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Growth  
; TITLE OF INVENTION: Development and Uses Thereof  
; FILE REFERENCE: 70166USNP  
; CURRENT APPLICATION NUMBER: US/10/696,616  
; CURRENT FILING DATE: 2003-10-29  
; PRIOR APPLICATION NUMBER: US 60/423,519  
; PRIOR FILING DATE: 2002-11-04  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 436  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-696-616-28

Query Match 66.1%; Score 1454; DB 16; Length 436;  
Best Local Similarity 67.7%; Pred. No. 2e-135;  
Matches 268; Conservative 65; Mismatches 63; Indels 0; Gaps 0;  
QY 31 ASSSLRFAVHRCRTARLEVKASGNTFGNYFQVATYGESHGCGVGVISGCPRIPLTEAD 90  
DB 30 SSPAVQISLRTQTRKNFOIQATGSSYTHFRVSTFGSHGGVGCIIIDGCPRIPLTEAD 89  
QY 91 LQVELDRRPPGQSRTSTKTDTCILSGTHEGVTGTPILVIVPNTDQIGSDHREIAN 150  
DB 90 LQVELDRRPPGQSRTSTKTDTCILSGTHEGVTGTPILVIVPNTDQIGSDHREIAN 149  
QY 151 VYRSHADATYFKYGVRAVOGGSSGRKTVGRVAAGALPKKILKLCGLLEILSVSKV 210  
DB 150 AYRPSHADATYDMKYGVRSVOGGSSARETIGRVAPGALAKILKQFAGTEILAYVSQV 209  
QY 211 HOVLPEDAVDYGSTVLEQIESNIVRCPPPEYAEKMDAIDRVVRGDSVGVITCVARN 270  
DB 210 HHVLPBELVDHENUTLEQIENNIVRCPPPEYAEKMDAIDRVVRGDSVGVITCVARN 269  
QY 271 VPRGLGSPVFDKLESELAKAMLSIPASNGFBIIGSGFAGTDLTGSEHNDEFYMDKAGSVRT 330  
DB 270 APRGLGTPVFDKLESELAKAMLSIPATKGFQFGSGFAGTFELTGLEHNDEFYMDKAGSVRT 329  
QY 331 RTNRSVGVGGISNVEIHFVKAFKPTPSIGVKONTVSRERQNVVELLARGHDPVAPRA 390  
DB 330 RTNRSVGGIQQGINSGEIINMRVAFKPTSTIGRKONTVTRDKVETEMIAARGHDPVAPRA 389  
QY 391 VPVBSMAALVMDQLMAHVAOCMEFALNTALQEPV 426  
DB 390 VPVBSMAALVMDQLMAHVAOCMEFALNTALQEPV 425

RESULT 15  
US-10-677-179-12  
; Sequence 12, Application US/10677179  
; Publication No. US20040082050A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/10/677,179  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US/09/743,207  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-677-179-12

```
Query Match      57.1%; Score 1257; DB 15; Length 307;
Best Local Similarity 81.9%; Pred. NO. 4.8e-116;
Matches 245; Conservative 22; Mismatches 26; Indels 6; Gaps 2;

QY 15 ARLAPRAIG---ALLEPAPASSSLRFAVHRCRTARLEVKASGNTFGNYFOVATYGESHG 70
Db 11 ARVLPARGGGGFRAFPESAPA--SLRFSVGRRRARLEVKASANVFGNYFOVATYGESHG 68

QY 71 GGVGCVISGCPRIPLTADLQVELDRRRPQOSRITSTRKETDTCKILSGTHEGVTTGTP 130
Db 69 GGVGCVISGCPRIPLTEADMQVELDRRRPQOSRITTPRKETDTCKILSGTHEGVTTGTP 128

QY 131 ILVIVPNTDOIGSDHREIANVYRPSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGAL 190
Db 129 IHVFVPNTDQRGDYSEMAKAYRPSHADATYDFKYGVRVAVQGGSSARETIGRVAAGAL 188

QY 191 PKKILKLCGLEILISFVSKVHVLPEDAVDYGSVTLQIESNIVRCPPDEYAEKMDAI 250
Db 189 AKKILKLSGVEILAFVSKVHVLPEDAVDYDTVMEQIESNIVRCPPDEYAEKMDAL 248

QY 251 DRVRVRGDSVGVVITCVARNVPRGLGSPVFDKLSSELAKAMLSIPASNGFEIGSGFAGT 309
Db 249 DKVRVRGDSIGGVVTCIARNVPRGIGSPVFDKLEAEALAKAMLSLPASKGFEIGSGFVFT 307
```

Search completed: August 25, 2005, 19:35:16  
Job time : 163 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 27, 2005, 02:58:29 ; Search time 4713 Seconds  
(without alignments)  
4410.628 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTVPKPOQVAHSRLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US10677179/runat\_25082005\_150240\_24835/app query.fasta\_1.583  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10677179 @CNG 1.1 5600 @runat\_25082005\_150240\_24835 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

denEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*

RESULT 1  
AR432943  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

AR432943  
Sequence 7 from patent US 6653531.  
AR432943  
AR432943.1 GI:40195584  
Unknown.  
Unclassified.  
1 (bases 1 to 1626)  
Cahoon,R.E. and Falco,S.C.  
Chorismate synthase from plants  
Patent: US 6653531-A 7 25-NOV-2003;  
Location/Qualifiers  
1. 1626  
/organism="unknown"  
/mol\_type="genomic DNA"

linear DNA 1626 bp PAT 18-DEC-2003

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2200	100.0	1626	AR432943
2	1800	81.8	1607	AK099850
3	1797.5	81.7	1635	AR432940
4	1541	70.0	1802	LECHOSYNB

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Alignment Scores:  
Pred. No.: 1.47e-138  
Score: 2200.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 6  
Length: 1626  
Matches: 429  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

5 1527 69.4 1701 8 LECHOSYNA  
6 1486.5 67.6 1754 8 CSCHOSYN  
7 1454 66.1 1311 6 CQ805346  
8 1454 66.1 1311 8 BT015785  
9 1454 66.1 1603 8 AY057519  
10 1257 57.1 966 6 AR432945  
11 1210.5 55.0 195583 8 AC134234  
12 1195 54.3 110000 2 AP006487\_1  
13 1192.5 54.2 339650 1 AP003583  
14 1191.5 54.2 1884 1 SSRPAROC  
15 1191.5 54.2 135551 1 D90906  
16 1176.5 53.5 299350 1 AP005370  
17 1066 48.5 1015 6 AR432942  
18 1066 48.5 1057 8 BT009607  
19 1047 47.6 349742 1 BX572090  
20 1038 47.2 348764 1 BX569689  
21 1025.5 46.6 306803 1 AB017161  
22 1010 45.9 348071 1 BX572100  
23 954.5 43.4 18062 8 AC084242  
24 950.5 43.2 110000 1 CR522870\_34  
25 946.5 43.0 110000 8 CR382125\_08  
26 944 42.9 110000 8 CR380951\_4  
27 943 42.9 110000 8 AB016817\_01  
28 942 42.8 110000 8 CR382137\_03  
29 938 42.6 1876 8 SCVGL148W  
30 938 42.6 3036 8 SCAROA  
31 938 42.6 22803 8 SCV11228  
32 930.5 42.3 110000 8 CR382130\_22  
33 930 42.3 1474 8 NCU25818  
34 928.5 42.2 10157 1 AB011207  
35 928.5 42.2 300693 1 AB017287  
36 889.5 40.4 110000 1 AP006841\_43  
37 887 40.3 659 8 AB122084  
38 876 39.8 59412 8 NCB7721  
39 874.5 39.8 301731 1 AB016934  
40 870 39.5 36677 2 AY714863  
41 864 39.3 299180 1 AP005080  
42 864 39.3 348526 1 CR378672  
43 862 39.2 10923 1 AB013761  
44 862 39.2 110000 1 BX936398\_31  
45 862 39.2 258050 1 AJ414153

## ALIGNMENTS

US-10-677-179-8 (1-429) x AR432943 (1-1626)

Qy 1 MetThrThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20  
Db 79 ATGACACCGTGGCCCAAGCCACAGACAGGTGGCGCACTCACGGGACCGCTCGCACCCGCG 138  
Qy 21 AlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40  
Db 139 GCGATCGGCGCCCTTCTGAGTTTCCCGCAGCCTCTCTCCCTCCGCTTCCGCGTGCAC 198  
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60  
Db 199 CGCTGCCGCACTGCTCGCTAGAGGTGAAGGCATCTGGAACACAGCTTTGGAAACTTACTTT 258  
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCys 80  
Db 259 CAGGTTGCAACCTATGGTGAATCTCATGGGGTGGTGTGGTGTGTATCAGTGGTTGT 318  
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100  
Db 319 CCACCTAGATTCCTACTCTAGGCGACACCTACAGTTGAAGTTCGATCGAAGACGGCCC 378  
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGly 120  
Db 379 GGACAGAGCAGAAATTAACCTCCACAAAGGAAGAGACTGATACATGCAAAATTTCTGTGAGG 438  
Qy 121 ThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThrAspGln 140  
Db 439 ACACATGAAGGGGTGACTACTGGAACGCCAATTCCTTATTTGTGCCAAACACACAGATCAA 498  
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAsnAlaThr 160  
Db 499 ATAGCAGTGATCACCGTGAATACCAATGTTATCCGACCTTCTATCGAGACCACT 558  
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180  
Db 559 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTCTCGGCAGAAAA 618  
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200  
Db 619 ACCGTTGGAAGGGTGGCTCGAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGA 678  
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220  
Db 679 TTAGAGATCTGTGCTGTTGTTTCCAAAGTGCAATCAGTTGTGCTCCCAAGACCGGTT 738  
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240  
Db 739 GATTATGGTCTGTAACTTTGGAACAGATAGAGACCAATCGTTAGATGCTCTGATCCA 798  
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgLysAspSerVal 260  
Db 799 GAGTACGACAGAGAAGATGATAGACCAATCGACAGATACGAGTTCGAGGGGATTCGGTC 858  
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280  
Db 859 GGTGAGTGATGATCATGTGCTGCTAGAAACGTTCTTCGGGGCTCGGTTCTCTGTGTTTC 918  
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300  
Db 919 GACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCTTCGCGAGCAACGGGTTTC 978  
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320  
Db 979 GAGATTCGACGGGATTCGCGGGACCGACTTTCGACAGGAAGTGAACATAATGATGAGTTT 1038  
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340  
Db 1039 TATATGGATTAAGCTTGGAAAGTGTGAGCACCGGACTAATCGCTCGGTTGGTGTGACGGGA 1098  
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360  
Db 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAAAGTTGCTTTTAAAGCCGACACCATCTATC 1158

Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380  
Db 1159 GGGGTGAAACAGAACACCGGTCAAGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGG 1218  
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400  
Db 1219 CGCATGACCATGCGTCCGCTCGAGCTGTTCTGTGTGGATCATGCGCGCTGTG 1278  
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420  
Db 1279 GTCTCTATGACACGAGTGTGGCGCACGTGGCTCAGTCGAGATGTTTCGGCTCAATACT 1338  
Qy 421 AlaLeuGlnGluProValGlySerPhe 429  
Db 1339 GCATTCAGAACAGTTGGCTCTTTC 1365  
RESULT 2  
AK099850 1607 bp mRNA linear PLN 24-JUL-2003  
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013110C10, full  
DEFINITION insert sequence.  
ACCESSION AK099850  
VERSION AK099850.1 GI:32985059  
KEYWORDS FLI CDNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
REFERENCE 1  
AUTHORS The Rice Full-length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-length cDNA Project Team,  
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,  
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,  
Kojima,K., Namiki,T., Ohmeda,E., Yahagi,W., Suzuki,K., Li,C.,  
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group, Otono,Y., Murakami,K.,  
Iida,Y., Sugeno,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
Kuroski,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,  
Ikedawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,  
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,  
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,  
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Otsu,N., Oca,Y.,  
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Yoshino,M. and Hayashizaki,Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
JOURNAL MEDLINE  
PUBMED 22752273  
REFERENCE 12869764  
AUTHORS 2 (bases 1 to 1607)  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,  
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,  
Hori,F., Hotta,I., Iida,Y., Ikeda,R., Inamura,K.,  
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,  
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,  
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,  
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,  
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,  
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,  
Namiki,T., Nariikawa,R., Niikura,J., Nishi,K., Nomura,K.,  
Numasaki,R., Ohmeda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,  
Otsu,N., Oca,Y., Otono,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,  
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,  
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,  
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,  
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,  
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,

**COMMENT**

## FEATURES

## ORIGIN

Alignment Scores:	
Pred. No.:	9.99e-112
Score:	.1800.00
Percent Similarity:	90.91%
Nearest Local Similarity:	83.49%
Query Match:	81.82%
DB:	8
Length:	1607
Matches:	349
Conservative:	31
Mismatch:	32
Indels:	6
Gaps:	2
JS-10-677-179-8 (1-429) x AK099850 (1-1607)	

15	AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro	30
89	GCDCGCGTCTCCGCCCGCGCGGGGTTCGCGCCTTCCGCGAGTCCGCGCG	148
31	AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValLys	50
149	GCT-----TCCCTCCGCTTCTCGTTCGCGCGCGCGCGCGCTCGCTAGAGGTGAAG	202
51	AlaSerGlyAenThrPheGlyAenTyrPheGlnValAlaThrTyrGlyGluSerHisGly	70
203	GCCTCTCAAAATGATTGTGGAACTACTTCCAGGTTGCACACTTTCGAGAGTCTCATGA	262
71	GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp	90
263	GGCGGGTGTGGTTCGCTGAATCATGGATGCCACCAGAAATCCCACTTACTTAAACGAGAT	322
91	LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys	110









```

Fumariaceae; Corydalis.
1 (bases 1 to 1754)
Schaller, A., Schmid, J., Leibinger, U. and Amrhein, N.
Molecular cloning and analysis of a cDNA coding for chorismate
synthase from the higher plant Corydalis sempervirens Pers
J. Biol. Chem. 266 (32), 21434-21438 (1991)
92042037
PUBMED 1718979
2 (bases 1 to 1754)
Amrhein, N.
Direct Submission
Submitted (18-JUL-1991) N. Amrhein, Swiss Federal Institute of
Technology, Institute of Plant Sciences, Sonnegstr 5, Zuerich
8092, Switzerland
Location/Qualifiers
1..1754
/organism="Corydalis sempervirens"
/mol_type="mRNA"
/db_xref="taxon:3464"
1..1754
/evidence=experimental
165..1508
/EC_number="4.6.1.4"
/codon_start=1
/product="chorismate synthase"
/protein_id="CAA43034.1"
/db_xref="GI:18256"
/db_xref="GOA:P27793"
/db_xref="UniProt/Swiss-Prot:P27793"
/translation="MASSLSTKPLSGRRRTTDSGWSYFQTSDLRLQNSQVQIS
VRRQAPLKVVAQSGSFGVQVTTGESHGSGVGVIDGCPFRPFSEADIQSDL
DRRTPQQRITTPRKETDKIVGVADFTTGSPIHISVPNTDORNDYSEMAKAYR
PSHADATYDFKVSQVSGSSARETIGRVAAGALAKKILKAVAGTEVLAVUSQAH
KVPLGVLVDHETLSLEQIESNIVRCPSDEYAEKMLAIDAVRVKGSVGVVTCMR
NVPRLGSVPFQKLEALAKAMSPATGFEFGSGFTFLTSEHNDFFYDNGR
IYTRNRSQGIQGGISNGEIIINRIAFKPTSTIGKQNTVTREREIEILARGRDPC
VYPRAPVMEANVALVLLDQLMLOAQGNLFSINPALQELPSETVSSAAASIQGV"
transit_peptide 165..335
mat_peptide 336..1505
/product="chorismate synthase"
/EC_number="4.6.1.4"

ORIGIN
Alignment Scores:
Pred. No.: 1 21e-90 Length: 1754
Score: 1486.50 Matches: 280
Percent Similarity: 85.43% Conservative: 60
Best Local Similarity: 70.35% Mismatches: 55
Query Match: 67.57% Indels: 3
DB: 8 Gaps: 2

US-10-677-179-8 (1-429) x CSCHOSYN (1-1754)

QY 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAla-----ArgLeuGlu 48
Db 273 TCDAAATCAATCGTCAGATCTCTGTCGACGA---CAAAACGCCCCCTTGAACACTAGTT 329
QY 49 ValLysAlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTrpGlyGluSer 68
Db 330 GTACAGGCATCCGGAGTTCAVTTGGGAAAGTCTTCCAAAGTTACAAACATATGGAGATCT 389
QY 69 HisGlyGlyValGlyValGlyValIleSerGlyCysProArgIleProLeuThrGlu 88
Db 390 CACGAGGAGGTTGGTGTGTGTCATGTATGATGTCCTCCACGGTTCCCACTCCGAA 449
QY 89 AlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThr 108
Db 450 GCTGATATACATCAGATCTTCACAGAGGAGCGCAGGTACAGCGCATACTACCCCA 509
QY 109 ArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGly 128
Db 510 AGAAAAGAACTGATACATGCAAAATATATTTCTGGAGTTGCTGATGGATTTCACAACTGGG 569

```

```

QY 129 ThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIle 148
Db 570 TCTCAATTCACATCAGTGTGCTTACACTGATCAGCGAGGAATGATTCAGTGAATG 629
QY 149 AlaAsnValTrpArgProSerHisAlaAspAlaThrTrpAspPheLysTrpGlyValArg 168
Db 630 GCAAAAGCTTATAGGCTTCTCATGCTGATGCTACTATGACTTCAAGTATGGTGTGAGA 689
QY 169 AlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGly 188
Db 690 TCAGTGCAGGGTGGGGAAGATCTTTCAGCAAGAGAAACATTTGGGAGAGTTTCAGCTGGA 749
QY 189 AlaLeuProLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSer 208
Db 750 GCTCTTGCACAGAGATTTCTCAAGGCTTATGCAGGACTGAGGTTCTAGCATATGTCTCT 809
QY 209 LysValHisGlnValIleValProGluAspAlaValAspTrpGlySerValThrLeuGlu 228
Db 810 CAAGCCCAAAAGTTGTTCTTCAGAGGGTTTGGTTGATCATGAGACTCTATCACTCGAA 869
QY 229 GlnIleGluSerAsnIleValArgCysProAspProGluTrpTrpAlaGluLysMetIleAsp 248
Db 870 CAGATTGAAGTAACATCGTTAGTGCACAGATTCAGAAATACCCGAGAGATGATGCT 929
QY 249 AlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAla 268
Db 930 GCTATTGATGCTGTCGAGTGAAGGAGACTCTGTTGTTGGAGTTGTGACATGTTATTG 989
QY 269 ArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAla 288
Db 990 AGGAACGTCACAGTGGGCTTGGTTCCACAGTATTATTGATAACTTGAAGCTGAGCTTGC 1049
QY 289 LysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGly 308
Db 1050 AAAGCTGCTATGTCATTACCAACAAAGGGGTTTGAATTTGGGAGTGGATTTTCAGGC 1109
QY 309 ThrAspLeuThrGlySerGluHisAsnAspGluPheTrpMetAspLysAlaGlySerVal 328
Db 1110 ACATTTCTGACTGGTAGTGAACATATATGATGTTCTACACAGCAGAGATGGAAGATC 1169
QY 329 ArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleVal 348
Db 1170 AGGACTAGAAACAAATCGATCTGTTGGAATACAGGGTGTGTTATATCCACGAGAAATCATA 1229
QY 349 HisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSer 368
Db 1230 AACATGAGAAATAGCTTTCAAGCCGACATCTACAAATTTGGAAGAAAGCAGATACTGTGACC 1289
QY 369 ArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaPro 388
Db 1290 AGAGAACGTGAAGAGATTGAATCTATGACGTGTGTCGCCACGACCCCTTGTTGTACTT 1349
QY 389 ArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAla 408
Db 1350 CGAGCTGTTCCAAATGTTGGAAGCTATGTTAGCCCTGTTGCTCTCTGATCACTAGTGTCT 1409
QY 409 HisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
Db 1410 CAACATGCTCAGGCAACCTTTTCTCCATCAATCTGCTTTTACAGAAACCTTTG 1463

RESULT 7
CQ805346 1311 bp DNA linear PAT 10-MAY-2004
LOCUS Sequence 1757 from Patent WO2004/035798.
DEFINITION CQ805346
ACCESSION CQ805346
VERSION CQ805346.1 GI:47111307
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1

```







this work. Shinozaki, K. (RIKEN GSC) and Becker, J.R. (ssp/salk) contributed equally to this work as PIs.

## FEATURES

source

Location/Qualifiers

```
1..1603
  /organism="Arabidopsis thaliana"
  /mol_type="mRNA"
  /db_xref="taxon:3702"
  /chromosome="1"
  /clone="RAPL09-47-O14 (R14512)"
  /ecotype="Columbia"
  1..122
123..1433
  /note="putative chorismate synthase"
  /codon_start=1
  /product="At1g48850/T24P22_3"
  /protein_id="AA09759.1"
  /db_xref="GI:15982824"
  /translation="MASSSLTSKSLGSLKSLSSLPSELRLSSPAVQISLRITQTRK
  NFQIQTSSYGVTHPRVSTFGESHGVCIIIDGPPRIPLTESDLOFLDRRPPGOS
  RITTPRKETDTCRISSGVSEGMWTCGPVHVPNTDORGLDYSMSVAYRSHADATY
  DMKYGVRVQGGSSARETIGRVAPGALAKKILKQFAGTBILAYVSQVHHVVLPEEL
  VDHEMLTUEQIENNIIRCPNFEYAEKIMAAIDAVRTKNSVGGVVTCLVRNAPRLGT
  PVFDKLEALKAKMSLPATKPFSGFAGTFLTGLEHNFDEFYTDENGRIKRTNRS
  GGIQGIENGSIINRVAKPTSTIGRKONTVTRDKVETEMIAARGHDPCCVPRVPM
  VEAMVALVLVQLMAQYAQCHLFPINPELOPQLIQEQPNATAL"
1434..1603
```

## 3' UTR

## ORIGIN

## Alignment Scores:

Pred. No.:	1,648-88	Length:	1603
Score:	1454.00	Matches:	268
Percent Similarity:	84.09%	Conservative:	65
Best Local Similarity:	67.68%	Mismatches:	63
Query Match:	66.09%	Indels:	0
DB:	8	Gaps:	0

US-10-677-179-8 (1-429) x AY057519 (1-1603)

Qy	31	AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValys	50
Db	210	TCTTCTCCGCGGTTTCAGATCTCTCTCGGTACCCAAACAGGAGAACTTCCAGATACA	269
Qy	51	AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyThrGlyGluSerHisGly	70
Db	270	GCTACTCGGAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTTGGAGATCATATGGA	329
Qy	71	GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp	90
Db	330	GGAGGAGTTGGTTGATCATTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	389
Qy	91	LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglys	110
Db	390	TTGCAATTTCGATCTCGATAGAGGAGGCTGTCAGAGCAGGATCACAACCTCTAGAAAA	449
Qy	111	GluThrAspThrCysGlyIleLeuSerGlyThrHisGlyGlyValThrThrGlyThrPro	130
Db	450	GAGACTGATCTTGGCGGATATCTCTGAGGTCCTGAGGATGTCAGCAGGACACCT	509
Qy	131	IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150
Db	510	ATCCATGTGTTTGTACCAACAACAGATCAGAGAGGACTTGATTACAGTGAATGTCGGTT	569
Qy	151	ValTyArgProSerHisAlaAspAlaThrTyArgPheTyThrGlyValArgAlaVal	170
Db	570	GCCTATAGACCATCGCATGCTGATGCACTATATGATGATGATGATGATGATGATGATGATG	629
Qy	171	GlnGlyGlyGlyArgSerSerGlyArgIleValGlyArgValAlaAlaGlyAlaLeu	190
Db	630	CAGGCTGAGGAGGAGGATCTTCACTAGAGACCATTTGGAAGAGTTGCTCTCGGACCTTG	689
Qy	191	ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal	210
Db	690	GCCAGAGAAATTTTGAAGCAATTTTCAGGAACTGAGATTTCTTGCCCTATATGTTCTCGCA	749

Qy	211	HisGlnValValLeuProGluAspAlaValAspTyArgSerValThrLeuGluGlnIle	230
Db	750	CACCATGTTGTACTTCCAGAGAAATTCGTAGACGACGAGAAATTTAACACTCGAACACATA	809
Qy	231	GluSerAsnIleValArgCysProAspProGluTyAlaGluLysMetIleAspAlaIle	250
Db	810	GAATAATTAACATTGTTCAGATGCCCTAATCCCGAGTATCGCGAAAGATGATAGCTGCGATT	869
Qy	251	AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn	270
Db	870	GATGCTGTACAGGACAAAGGGAACCTCTGTTGGTGTGTGTGACCTGATTTGTCGGAAT	929
Qy	271	ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla	290
Db	930	GCTCCAGTGGGCTTGTGATCACCCTGTTTCGATAAATTTTGAAGCAGAACTGGCAAAAGCT	989
Qy	291	MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp	310
Db	990	TGATGTCGTACTCTGCACTAAGGATTTGAGTTTGAAGCGGCTTTCGAGGTACCTTT	1049
Qy	311	LeuThrGlySerGluHisAsnAspGluPheTyThrMetAspLysAlaGlySerValArgThr	330
Db	1050	TTGACTGCTCTTGAACACAATGATGATCTCTATACCGATGAAATGGAAGAATACGTACC	1109
Qy	331	ArgThrAsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPhe	350
Db	1110	AGAACCACCGATCTGTTGGAATTCAGGAGGATCTCAATGTTGTAATTAATAAATG	1169
Qy	351	LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu	370
Db	1170	AGATAGCTTCAAGCCAAACATCAACAATTTGGAAGGAGCAAAATACGGTAACACAGAGAC	1229
Qy	371	ArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAla	390
Db	1230	AAGGTAGAAACCGAAATGATGTCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1289
Qy	391	ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal	410
Db	1290	GTGCCAATGTGGAGCAATGGTGGCTCTAGTCTTGTGGATCAATTTGATGGCGCAATAC	1349
Qy	411	AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal	426
Db	1350	GCACAATGCCATTTGTTTCCATAATAATCCAGAGTTGCAGGAACCTCTC	1397
RESULT 10			
AR432945			
LOCUS			
DEFINITION			
Sequence 11 from patent US 6653531.			
ACCESSION			
AR432945			
VERSION			
AR432945.1			
KEYWORDS			
SOURCE			
Unknown.			
ORGANISM			
Unclassified.			
REFERENCE			
1 (bases 1 to 966)			
AUTHORS			
Cahoon, R.E. and Falco, S.C.			
TITLE			
Chorismate synthase from plants			
JOURNAL			
Patent: US 6653531-A 11 25-NOV-2003;			
FEATURES			
Location/Qualifiers			
source			
1..966			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Alignment Scores:			
Pred. No.:			
1,426-75			
Length:			
966			
Score:			
1257.00			
Matches:			
245			
Percent Similarity:			
89.30%			
Conservative:			
22			
Best Local Similarity:			
81.94%			
Mismatches:			
26			
Query Match:			
57.14%			
Indels:			
6			
Gaps:			
2			
DB:			





```
QY      97 ----- 97
Db      83229 GTCTCTTGAAAAATCCATTATTTCTGTCTGCACAGCTGGCTTCTGTCAAAACAGGATTA 83170
QY      97 ----- 97
Db      83169 TGCAAATGCAATCCTGCTAACCTCTTGCGAATATGATTTGTTATGTTATACACCATTA 83110
QY      98 -----
Db      83109 ATGTTAACTGACATAAAATTTTGAATTATCATCAGACGGCCAGCCAGACGAGATAAACAC 83050
QY      107 rThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyVal----- 125
Db      83049 CCCAAGAAGGAGACTGACACTTGCAAAATTTCTTTTCAGGACACATGAAGGTAT-GGTGT 82991
QY      125 ----- 125
Db      82990 CTTTCAGGACATGGGATAAACTTGGAAAAATATTTTGAATTTAGAGTATTAGCTTGTCT 82931
QY      126 -----
Db      82930 TCATATTCACGATAGCAATTGATGGGTTTCTGCAGGAATGACCACTGGGACACCAATTCA 82871
QY      132 uValIleValProAsnThrAspGlnIleGly----- 142
Db      82870 TGTTTTGTCCGGAACACAGATCAGAGAGGGGGTGTAAAGTACCTAAGTTTTTTTCGCATTA 82811
QY      142 ----- 142
Db      82810 ATGTATTTGAATACTACCTCCGTCCTCCAAAATATAACAACCTTTGGCTAGCAATCTGGACA 82751
QY      142 ----- 142
Db      82750 CACAATTTGCCAGATTCATAGCTAAATGCTTATATTTTGGGACAGAGGGGTAACATA 82691
QY      143 -----
Db      82690 TTAAGCTGGAATAACCATATTCATCCACAGGATTCACAGTGAATGGCTAAGGCCTA 82631
QY      152 rArgProSerHisAlaAspAlaThrAspPheLysTyrGlyValArgAlaValGln-- 171
Db      82630 CAGACCTTCACATGCAGATGCAACTTATGACTTCAATACGGTGTAGACGATGCAGGT 82571
QY      171 ----- 171
Db      82570 ACTTTTTTTCATGTTTATATGTTGTAATATAGCCATTCAATTGAACCATCAACTGTAGTA 82511
QY      171 ----- 171
Db      82510 CCTAAACTTCTCTTATTTGAATGTCTAAACCTGCTGCTACCTAAACCTTCTTACTTGAAT 82451
QY      172 -----
Db      82450 GCCTAACCAACATCTTCTATCAACCGAATGATAGGAGGTGGAAGATCATCAGCAAGAGA 82391
QY      180 sThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysCysG1 200
Db      82390 GACCATTGGAGGGTGGCTGCAGGAGCTCTTGCAAAAGAAAATTTCTTAAGCTCAATCTGG 82331
QY      200 Y----- 200
Db      82330 AGTAGAGGTGTGATCCTACTTGGCAACCTTTTACTTCTTAACGTATCAAGATTTGATCAA 82271
QY      200 ----- 200
Db      82270 GCATCGCAATTGCATGTTTACTTATTCATGACAACTTGGACGCTGAAGAAAAATGAATC 82211
QY      200 ----- 200
Db      82210 AGCATTCGAATATACATCCTAAATATATGCTGAAGCCATGCAAAAATTTCTGTCCAATGTCT 82151
QY      201 -----
Db      82150 CAFAGCAGCAAAATATTTTCAAATTGCGAGTCTTGGCGTGTGTGTCCAAGGTGCATCAAGT 82091
QY      213 lValLeuProGluAspAlaValaAspTyrGlySerValThrLeuGlu----- 228
Db      82090 TGTACTACCAAGAGATGCCGTTGATTATGACACTGTGTAACATGGAACAGGTCTCTGGTTC 82031
QY      228 ----- 228
Db      82030 CCACCATTTTCAAGTTTAAACACTTCTGGACAGTCTCTGAAGGCATTTTGTCAATATTATA 81971
QY      229 -----
Db      81970 TTTATTTTATTTTCCAGATAGAAGCAACATTTGTAGATGTCCTGATCCGAATATGCAC 81911
QY      244 lLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValI 264
Db      81910 AGAAGATGATTGATGCAATCGAATAAGTACGAGTTAGAGGTGATTCCGATTGGTGGTGG 81851
QY      264 leThrCysValAlaArgAsnValProArg----- 273
Db      81850 TCACATGCATTTGCAAGAAATGTTCTCGTGTAAAGTATTTAAACCTGCAGTTATGTATTAT 81791
QY      273 ----- 273
Db      81790 TTTGATATATGTCCTCCAGCTAAATAAATAAATAAGTCATCTGTTTGCACCTCTCTGCT 81731
QY      273 ----- 273
Db      81730 TATGACCGCCTGCTTTTGGAAATTTGCACAAAATCTGTTTGAACATTTGAAGCTTGATCAT 81671
QY      273 ----- 273
Db      81670 CATTAGAAAAATATTTTCTCTGCTGAGCTACTGATATATGATAGACAGCAACAATTTGT 81611
QY      274 --GlyLeuGlySerProValPheAspLysLeuGlySerGluLeuAlaLysAlaMetLeuS 293
Db      81610 AGGGGATGGCTCTCTGTTATTTGACAACCTTGAGGCTGAATTTGGCGAAAGCTATGCTTT 81551
QY      293 erIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAla----- 307
Db      81550 CTCTCTCTGCAAGCAAGGGGTTTGAGATCGGAGTGGATTTCGAGGTAAATTGAATTAATC 81491
QY      307 ----- 307
Db      81490 TGCACATAATTAGTTGATTATTTTTCAAATAGCTGTGCACTGCCAACTACTTTTGCTAC 81431
QY      308 -----
Db      81430 TGGCCTCAACTACTAATAATACTTGTGCCAGGTACTGACTACACTGGAAGTGAGGATA 81371
QY      317 snAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyG 337
Db      81370 ATGATGATTTCTATATGATGAGGCTCGAAATGTGAGACAACAAGAACTAAATCGTTCCAGCG 81311
QY      337 lyVal----- 338
Db      81310 GTGT-GCAGGTTTATCTCTAGTCCCTGTTGTTTACAGTTATATTATTTTATTTGTTTCG 81252
QY      338 ----- 338
Db      81251 CTACATCTAGCTGCAATCTAGATTAAAAAATACTAGTAATAAGTATATAACACCGTTGTCT 81192
QY      339 -----
Db      81191 TAGTATTTTATAATGTGTATGTCCATGCGAGGAGGATATCAATGGTGAAATTATATACT 81132
QY      350 heLysValAlaPheLysProThrProSerIleGlyVal----- 362
Db      81131 TCAAAGTAGCTTTTCAAGCCCAACCGCGACTATTGGGGT-GAGATATTCGGGTCAAGTAAC 81073
QY      362 ----- 362
```



Qy 372 nAsnValGluLeuLeuAlaAaGlyArgHisProCysValAlaProArgAlaValPr 392  
Db 82989 GCAAGTAGCTCTTCGGGAGAGGAGACAGATCCCTGTGTGCTCTGCTGTCC 82930  
Qy 392 oValValGluSerMetAlaLaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGl 412  
Db 82929 GATGCTAGAGCATGGTGGCGCTGCTTCTGTTGATCACTGCTTCAACAGATTGCCA 82870

Qy 412 nCysGluMetPheAlaLeuAsn---ThrAlaLeuGlnGluProValGly 427  
Db 82869 GTGTTCCGCTGCTTCATTGATCCATCAACGGCACCGAACCTCTGGGA 82821

RESULT 13  
AP003583/c 339650 bp DNA linear BCT 28-NOV-2001  
LOCUS Nostoc sp. PCC 7120 DNA, complete genome, section 3/19.  
ACCESSION AP003583 BA000019  
VERSION AP003583.1 GI:17129939  
KEYWORDS  
SOURCE Nostoc sp. PCC 7120  
ORGANISM Nostoc sp. PCC 7120  
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

REFERENCE 1  
AUTHORS Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S.,  
Watanabe,A., Iriiguchi,M., Ishikawa,A., Kawashima,K., Kimura,T.,  
Kishida,Y., Kohaza,M., Matsumoto,M., Matsuno,A., Muraki,A.,  
Nakazaki,N., Shimpoto,S., Sugimoto,M., Takazawa,M., Yamada,M.,  
Yasuda,M. and Tabata,S.  
Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium Anabaena sp. strain PCC 7120  
DNA Res. 8 (5), 205-213 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
11759840  
2 (bases 1 to 339650)  
Kaneko,T.  
Direct Submission  
Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research  
Institute, The First Laboratory for Plant Gene Research; Yana  
1532-1, Kisarazu, Chiba 292-0812, Japan  
(E-mail: kaneko@kazusa.or.jp,  
URL: http://www.kazusa.or.jp/cyanobase/,  
Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)

FEATURES  
source  
1..339650  
/organism="Nostoc sp. PCC 7120"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:103690"  
/note="synonym:Anabaena sp. PCC7120"  
complement(98..2698)  
/gene="al10594"  
complement(98..2698)  
/gene="al10594"  
/note="ORF ID:al10594"  
/codon\_start=1  
/transl\_table=11  
/product="aminopeptidase"  
/protein\_id="BAB72552.1"  
/db\_xref="GI:17129940"  
/translation="MPHSYFDQDSNAHKSFELPGARPHYPNDRPGQVEHIFLDNLND  
IPQSYEGSCIRLLPIRNGIDRLTLDVNLQIRSVQVDEVGQDFDYDGEOLVQLSQ  
PTEIGKRLIAAYAVEKPGRGYFIQPSHYPNKTVOVMTQGEDEDSRFWPCFDYP  
GOLSTSEIRVRVPKSLMAISNGELDIOEDGKHLYHWSQQIHPYTLMTLAVADPAE  
IRDEWGRPVYVKEGREADAKSMGKTPTMIEFLSKYGYAYAPKPAQVCVSDFI  
FGGMENTSTLLTDCLDERAALDNTELSLVHELHAFQGLDLVKKHWSHAWIKE  
KNASYSYVMEQEYSAEAAAYRLLEARSYSDSGSYRRPMTVHYVREAIELYDRH  
TYEGSCYHMIARELGBDLFWAIIQTFRONAHKTIVETIDLLRAIEKATGNLAFLE  
DOYVYRGDPDVKESYVWGEAKLAKVTVTOVPADKDKVDFLRIPIGFGYNOTEA  
AAKLTQFTVRPHVERQSFPPLTEKPDFISIEDVGNHKLTVSLYPIPELKAQASDP  
DPISRYYAALAKKGLLEAYKALAASLKNPFGVGRVYEVAKLAEIKLDAQFOGLVE  
GLADKNAVRRVAQALAKITLASYKAVGVVKGDDGASYESASACALGAIARINPF  
DDKSKBEAKIKLQSLQVEERAGNVEVRSGAIGGLAEIKTBEAALDILLEYTKLGVPQ  
FLRLAAIRALGKISPGQSPVNLRIIARLGELSKENFFILMAVVAALQWETPKAMG  
VLSRLADQTADGRVFRFAEEELANVQKNIGTIDNALRLREBFDQLKQONQLKRLLEN

gene  
CDS

gene  
CDS

gene  
CDS

gene  
CDS

gene  
CDS

LEAKSOTVKQ"  
complement(2862..3059)  
/gene="as10595"  
complement(2862..3059)  
/gene="as10595"  
/note="ORF ID:as10595"  
similar to cytochrome d oxidase chain I"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB72553.1"  
/db\_xref="GI:17129941"  
/translation="MRHCVFHILLSPLYLKIIILNIQWKPSDRPHVMGLIYYAFRVM  
IAIGFFLAGLMLVSTLFFEVH"  
complement(3151..3810)  
/gene="all0596"  
complement(3151..3810)  
/gene="all0596"  
/note="ORF ID:all0596"  
unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB72554.1"  
/db\_xref="GI:17129942"  
/translation="MGVCHSKDVSEFCLEGRPFMEFVIKDGKYLKGLLLTFADGECYVK  
LSQLRFTFDWRLPTGTLQVGEKKYDAKTGKVTLKAEAVMAARTETSSAVKQVPAM  
ENTKPKPEKAKATILVCQKSDCKMRGKALCOALEATLSDRGLEDOVTIKGTGCMKNC  
KAGPNLWMPDKTRYTRIADQVFKLWKNHFSQSDSHSHQAVISINTKFEPPTRVSA  
F"  
complement(3877..4134)  
/gene="as10597"  
complement(3877..4134)  
/gene="as10597"  
/note="ORF ID:as10597"  
hypothetical protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB72555.1"  
/db\_xref="GI:17129943"  
/translation="MKQFSSEVEKSVVEVDWADRMQVYORLOLEIIPCWCNEANQPLK  
VELVTPMTVVQLWSVMRQFTSSRQDLIYSLEINWQSRYPIS"  
4451..5146  
/gene="alr0598"  
4451..5146  
/gene="alr0598"  
/note="ORF ID:alr0598"  
unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB72556.1"  
/db\_xref="GI:17129944"  
/translation="MYTMRPLLSTLSLVLLQGFVPIVQAQSGENTGDVASTAIQOV  
IAANMTNSADGNFYPERLMNRASILVKAFRLNKQEAANQPKSITVADVPKSHWA  
YQDIQTVLKTDVMKRYGNLFFPNQVRTRABALAIFAQYGVGFQSDVVNEILAPYP  
DANSIPNWARRAIATVITEGFIINTDAQNNITPLRPMTRGDMAYLLSKYLQKQKQPD  
PVVPTTDSQPLP"  
5596..7503  
/gene="alr0599"  
5596..7503  
/gene="alr0599"  
/note="ORF ID:alr0599"  
/codon\_start=1  
/transl\_table=11  
/product="1-deoxy-xylulose 5-phosphate synthase"  
/protein\_id="BAB72557.1"  
/db\_xref="GI:17129945"  
/translation="MHLSEITHPNQLHGLSIRLQQLIARQIRDKHQLTVAFEGGHLGP  
GLGVNLTGLYQTLTDLRDKVINDVGHQYAPHKLIITGRYSDFHTLRQKDGIAYLKR  
GENKDFHFGHASTSI SAALGALARDMNGEKFAVALIGDGLATGCMALAEANHAG  
HLPNLGLLVNLNENSI SPNVGAI PLYLNKRLSPVQRTKONFEQFKHI PVFGES  
LSPBLGRIKGMKRLAVPKVGAFFELGFTYMGPDVGNHEELIATFQQAHQIAGPVL  
VHVATIKGKGYELAEKQVGHQATPPNLTGKAIPSNKPKPAPPAIKVFSHTLVKLAE  
QNPKIIGITAMATGTCLQAKLPNQYIDVGAEGHATVLAAGLATGEGMRPVAAIY  
STFLQRAYDQIIHDVCTIQLNLPVFFCLDRAGIVGSDGPTHQGMVDIAYLRCTIPNIVMA

PKDAEMQRMVVTGIEHTTGPAMRFRNGYGVPLMEEGWBPLEIGKGEILRNGDDV  
LIIGYTWYPSQAAEILSEHGEIATVINARFVKPLDTELIVPLAOKIGRVVTLLEG  
CWGGFGSAVALADLADVVPVKRIGIPDLVVEHATPDSEKAEGLTSRQIAERVLQ  
AVFQKQPSVS"

gene

/gene="alr0600"

CDS

/db\_xref="GI:17129946"

/note="ORF ID:alr0600"

unknown protein"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB72558.1"

/db\_xref="GI:17129946"

/translation="MADELDKVKIECFILSKLHSMKLNFSMKLNLGRWIATTVFCL  
SAIAFWQCAFADTSAMADNSTYLIASTDATSLRNKSKFIRDTOKNVKDAATNA  
SRVDDATDNGSTAERKANDASRIKRAEEDAAETAVDNNLNINVERTIDNIGRAFT  
NN"

gene

complement (8399..9469)

CDS

/gene="alr0601"

complement (8399..9469)

/gene="alr0601"

/note="ORF ID:alr0601"

hypothetical protein"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB72559.1"

/db\_xref="GI:17129947"

gene

complement (9680..10618)

CDS

/gene="ntcb"

complement (9680..10618)

/gene="ntcb"

/note="ORF ID:alr0602"

ntcb gene product"

/codon\_start=1

/transl\_table=11

/product="nitrogen assimilation transcriptional activator"

/protein\_id="BAB72560.1"

/db\_xref="GI:17129948"

gene

complement (10861..11166)

CDS

/gene="alr0603"

complement (10861..11166)

/gene="alr0603"

/note="ORF ID:alr0603"

putative transcriptional regulator"

/codon\_start=1

/transl\_table=11

/protein\_id="BAB72561.1"

/db\_xref="GI:17129949"

gene

complement (11163..11522)

CDS

/gene="alr0604"

complement (11163..11522)

/gene="alr0604"

Alignment Scores:

Pred. No.: 339650

Score: 1192.50

Percent Similarity: 70.82%

Length: 339650

Matches: 230

Conservative: 71

Best Local Similarity: 54.12% Mismatches: 87  
Query Match: 54.20% Indels: 37  
DB: 1 Gaps: 4

US-10-677-179-8 (1-429) x AP003583 (1-339650)

QY 25 LeuLeuGluPheAlaProAlaSer---SerSerLeuArgPheAlaVal----- 39

Db 232538 CTTCTTAATTTTACTCTCTGATGACCATGCTTCCTTCGTTTCCCATCACCAAAATATAT 232479

QY 40 -----HisArgCysArg----- 43

Db 232478 AGATAACATAAAATGTCAAGAAACTTCTTTTGCTCAATATTACGAAAAAAGTACAAGACACA 232419

QY 44 -----ThAlaArgLeuGluValValys 50

Db 232418 AAGATTAATTTTACGGAACCTTAGGTACAAATACATAAGCGCTCTCTTTTCTCTCAAG 232359

QY 51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70

Db 232358 ATCATGGGCAATACCTTTGGACATCTGTTTCGGATCACTACTTTTGGCGAGTCCACCGGT 232299

QY 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAse 90

Db 232298 GGAGCGGTGGGGTGTGATTGACGGCTGTCTCGCTACTCGAAATTTCTCTCGAGGAA 232239

QY 91 LeuGlnValGluLeuAspArgArgProGlnSerArgIleThrSerThrArgLys 110

Db 232238 ATTCACTGGGAGTTAGACCGCAGACGCCGAGACAGATAGATTACAAACGCTCGTAAG 232179

QY 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130

Db 232178 GAACGACACACTGTGAGATTTTCTCGGGTATATGAAGGCAAACTCTGGGAACGCCA 232119

QY 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150

Db 232118 ATATCAATTTTGGTCCGAATAAAGATACCTCGTCCCAAGATTATGATGAGATGGCGCAA 232059

QY 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170

Db 232058 AAGTATCGTCTCTCCCATGCTGATGCAACTTATGATGCTAAATATGGGATTCGTAAATGG 231999

QY 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190

Db 231998 CAAGTGGCGGTAGATCTCTCAGCGCGGAGACAAATGTTAGTAGTAGCCGAGTCCGATC 231939

QY 191 ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210

Db 231938 GCTAAAAAAATCTCCGCAAGTTGCTAATGTAGAAATTATAGGCTACGTCAACAGTATC 231879

QY 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230

Db 231878 AAAGACTTG-----GAAGGTGTTGTAGATCCCAATCCGTCACCTAGACCAAGTG 231828

QY 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250

Db 231827 GAAGCAATATCGTTCCTGCTGTCCTGATGGCGAATTTGGCAGACCGCATGATGAATATC 231768

QY 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270

Db 231767 GAACAAACAGGTAGACAAAGGTGATTCTATCGCGCGTGTAGTGAATGCGTAGCGCGAAT 231708

QY 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290

Db 231707 GTTCTTAAAGCTTGGGTGAACCTGTATTTCATAGTTGGAAGCAGATATCGCCCAAGCT 231648

QY 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310

Db 231647 GTGATGCTTTTACCTGTAGCAAAAGGTTTTGAATCGTTTTCGGGACTCTG 231588

QY 311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330

Db 231587 CTAACAGGATTTGAGCATACGACGAATATATATATGATGAAGAAACGGTGAGATTGCGACA 231528





QY 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252  
Db 780 AATATTGTTGTTGTTCCGATGAAGATGCGCAAAATGATTGACCCATTGACCAG 839  
QY 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnValPro 272  
Db 840 GTCTACGACAAAGAACTCATTTGGGGGGTGTGGAATGTGCGATGCTCCATACGCTCC 899  
QY 273 ArgGlyLeuGlySerProValPheAspLeuGlySerGluLeuLeuAlaValMetLeu 292  
Db 900 AAGGGTTGGGGAACAGGTTTGTATGATTGGAAGCGGATTTGGCAAGGCAATGATG 959  
QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312  
Db 960 TCCCTCCCGCTACCAAGGGTTTGTAGTTGGCTCTGGTTTGGCCGTACTTACTGACT 1019  
QY 313 GlySerGluHisAsnAspGluPheTyrMetAspLeuAlaGlySerValArgThrArgThr 332  
Db 1020 GSTAGTCAACATAATGATGATGACTACTCTGATGAAGCGGGGGAATGGCGTACTCGCAC 1079  
QY 333 AsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLysVal 352  
Db 1080 AACCGTTCTGGGGGGTTCAGGAGGAGCATCAGCAAGCGGCAACCATATTATTATGCGCAT 1139  
QY 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372  
Db 1140 GCATTTAAACCTACTGCCACCATTTGTCAGACCAAAACCTTAGTAACATTTGGGAG 1199  
QY 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392  
Db 1200 GAAACACCATCTAGCCGCTAAAGGTGCGCATGATCTCTGCTTTGGCCCGGGCTGTGCC 1259  
QY 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGluLeuMetAlaHisValAlaGln 412  
Db 1260 ATGTTGGAAGCGATGCTGATAGTACTTTGTGACCATTTGTGCGATTTCCAAAGCCAG 1319  
QY 413 CysGlu 414  
Db 1320 TGTAAA 1325  
RESULT 15  
D90906/c  
LOCUS  
DEFINITION Synchocystis sp. PCC 6803 DNA, complete genome, section: 8/27,  
920916-1056466  
ACCESSION D90906 AB011339 BA000022  
VERSION D90906.1 GI:1652492  
KEYWORDS  
SOURCE Synchocystis sp. PCC 6803  
ORGANISM Synchocystis sp. PCC 6803  
Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
REFERENCE  
1 Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Suzuki, T., Miyajima, N., Sugita, M., and Tabata, S.  
Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome  
DNA Res. 2 (4), 153-166 (1995)  
96127529  
PUBMED 8590279  
REFERENCE  
2 Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugita, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuoka, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S.  
Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions  
DNA Res. 3 (3), 109-136 (1996)  
97061201  
PUBMED 8905231  
REFERENCE  
3 (bases 1 to 135551)

AUTHORS  
TITLE  
JOURNAL

Tabata, S.  
Direct Submission  
Submitted (28-JUN-1996) Satoshi Tabata, Kazusa DNA Research  
Institute, The First Laboratory for Plant Gene Research, Yana  
1532-3, Kisarazu, Chiba 292-0812, Japan  
(E-mail: tabata@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/,  
Tel: 81-438-52-3933 (ex. 2330), Fax: 81-438-52-3934)  
Potential protein coding regions were assigned on the basis of  
similarity search of the ORFs and GeneMark analysis.

## COMMENT

FEATURES  
source

1. 135551  
/organism="Synchocystis sp. PCC 6803"  
/mol\_type="genomic DNA"  
/strain="PCC6803"  
/db\_xref="taxon:1148"  
/note="synonym: Synchocystis PCC6803"  
/complement(27..281)  
/gene="ss13712"  
/complement(27..281)  
/note="orf ID: ss13712"  
/unknown protein  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA17414.1"  
/db\_xref="GI:1652493"  
/translation="MGETKLADQRWQTHDCKDYLSYIELHRRYELEVEVVASFWP  
HVPPEKTYLLAWEXELIILKRSFPNKECWRWGWQPFGLKLS"  
/complement(634..1722)  
/gene="aroc"  
/complement(634..1722)  
/note="orf ID: sl11747"  
/codon\_start=1  
/transl\_table=11  
/product="choiximate synthase"  
/protein\_id="BAA17415.1"  
/db\_xref="GI:1652494"  
/translation="MGNTFGSLFRITTFBSHGSGVGVIIIGCPPLRISPEIIOVDL  
DRRPGOSKIITPRKADQCILSGFEGKLTGPIALIVRNKARDSDYNEMAYKIR  
PSHADATYAAKIRGNWQGRSARETIGRAAGAKI LAQPNFVYVAKSIQ  
DIATVNSNTVTLEQVNSNIVRCDECAERMIERIDQVLRKDSIGVWEICARNAP  
KGLGPEVDFKLEADLAKAMSLPATKPFEGSGAGTLLTGSHNDYVLYDEAGEMRT  
RTNSGGVGGISNGEPIIMELIAKPTATIGOEQKTVSNIGETTLAAGRHDPCLVP  
RAVPMVEMAAVLVCDHLIRFQAQCKTL"  
/complement(1929..2315)  
/gene="rpl12"  
/complement(1929..2315)  
/note="orf ID: sl11746"  
/codon\_start=1  
/transl\_table=11  
/product="50S ribosomal protein L12"  
/protein\_id="BAA17416.1"  
/db\_xref="GI:1652495"  
/translation="MSRATDQILEQLKSLSLLEASLVKQIEEAFGVSAAPVGGWM  
AAAAAFAAEAEKEITFDVILEVFPADKIAVLKVRITITGLKREAKELVESTPRAI  
KEATGKDDAEAIKKQIEAGGAAYK"  
/complement(2403..2924)  
/gene="rpl10"  
/complement(2403..2924)  
/note="orf ID: sl11745"  
/codon\_start=1  
/transl\_table=11  
/product="50S ribosomal protein L10"  
/protein\_id="BAA17417.1"  
/db\_xref="GI:1652496"  
/translation="MGRIENKATVSDVQLFQDAQMTVIDYQGLTVAEITDLNR  
LRPGGCTKATNTLVRRLAGQAWSPMEFLTGTAILVLKEDLGAIKAYKKFKQ  
DTKTELRGVLEKSLTQADVEAIGDLPSEQLMGQTAGGINALATKIALGIKEVPA  
SVARGLOAHVDKE"  
/complement(3171..3887)

## gene

## CDS

## gene

## CDS

## gene

## CDS

## gene





```
QY 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnValPro 272
Db 1128 GTGCTACGACAAAAGACCTCATTTGGGGGGTGGGATGTGGATCGTAACGCTCCC 1069
QY 273 ArgGlyLeuGlySerProValPheAppLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db 1068 AAGGGGTGGGGGAACCAAGTGTGTGATAGTTGGAAGCGGATTTGGCCAAGGCAATGATG 1009
QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 1008 TCCCTCCCCGCTACCAAGGGGTTTGAGTTTGGCTCTGGTCTTTCGCCGTACTTTACTGACT 949
QY 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 948 GGTAGTCAACATTAATGATGAGTACTACTCTGATGAAGCGGGGAATGGCGTACTCGCAC 889
QY 333 AsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLysVal 352
Db 888 AACCGTTCTGGGGGGTTCAGGGAGGCATCAGCAACGCCAACCATTATTATGCGCATATA 829
QY 353 AlaPheIleProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
Db 828 GCATTTAAACCTACTGCCACCATTTGGTCAAGACAAAACCGTTAGTAACATTGGGGAG 769
QY 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
Db 768 GAAACCACTTACCCGCTAAGGTCCCATGATCCCTGGCTTTGGCCCCGGCTGTGCC 709
QY 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db 708 ATGGTGAAGCGATGCTGCATTAGTACTTTGTGACCATTTTGTGGCATTTCCAAGGCCAG 649
QY 413 CysGlu 414
Db 648 TGTAAA 643
```

Search completed: August 27, 2005, 06:18:20  
Job time : 5002 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 27, 2005, 01:03:39 ; Search time 619 seconds  
(without alignments)

4102.696 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USFTO.spool\_p/US10677179/runat\_25082005\_150239\_24825/app\_query.fasta\_1.583  
-DB=N Geneseq 16Dec04 -QFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLCN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10677179 @CN\_1\_1\_708 @runat\_25082005\_150239\_24825 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	3 AAZ50438	Corn clon
2	1797.5	81.7	1635	3 AAZ50435	Corn clon
3	1461	66.4	1593	3 AAC40227	Arabidops
4	1454	66.1	1311	12 ADN73862	Thale cre
5	1454	66.1	1311	13 ADT91426	Arabidops

6	1257	57.1	966	3 AAZ50440	Aaz50440 Rice clon
7	1191.5	54.2	1089	13 ADS47934	Ads47934 Bacterial
8	1186.5	53.9	1086	13 ADT44117	Adt44117 Bacterial
9	1182.5	53.8	1086	13 ADT45459	Adt45459 Bacterial
10	1066	48.5	1015	3 AAZ50437	Aaz50437 Wheat cio
11	1035	47.0	1080	13 ADT46302	Adt46302 Bacterial
12	981	44.6	1068	13 ADS58432	Ads58432 Bacterial
13	938	42.6	1131	13 ADT47171	Adt47171 Bacterial
14	934	42.5	1137	12 ADP98700	Adp98700 C. albica
15	929.5	42.2	1239	13 ADR85489	Adr85489 Aspergill
16	888.5	40.4	1074	8 ACA24777	ACA24777 Prokaryot
17	869.5	39.5	1421	13 ADS49229	Ads49229 Bacterial
18	862	39.2	1086	8 ACA54193	ACA54193 Prokaryot
19	860	39.1	1086	4 AAS52498	Aas52498 E. coli D
20	860	39.1	1086	8 ACA32569	ACA32569 Prokaryot
21	860	39.1	1086	13 ADT48732	Adt48732 Bacterial
22	860	39.1	3100	8 ACC79525	Acc79525 Aroc gene
23	859	39.0	1134	11 ACH99580	Ach99580 Klebsiell
24	858.5	39.0	1071	12 ADO57719	Ado57719 Actinobac
25	858	39.0	1056	8 ACA36973	ACA36973 Prokaryot
26	857	39.0	1086	4 AAS53974	Aas53974 Klebsiell
27	855	38.9	1083	8 ACA35743	ACA35743 Prokaryot
28	850	38.6	1083	8 ACA45011	ACA45011 Prokaryot
29	850	38.6	1089	10 ADE99834	Ade99834 Bacterial
30	850	38.6	1363	13 ADR84902	Adr84902 Aspergill
31	850	38.6	7363	13 ADR84315	Adr84315 Aspergill
32	847	38.5	1239	4 RAD06952	Rad06952 aroc gene
33	847	38.5	1690	2 AAZ22883	Aaz22883 E. coli w
34	847	38.5	4530	4 RAD06958	Rad06958 pME4 comp
35	846.5	38.5	1074	8 ACA42924	ACA42924 Prokaryot
36	845.5	38.4	34980	13 ADT05648	Adt05648 Haemophil
37	845	38.4	110000	6 ABA92787_0	Ab92787 Buchnera
38	845	38.4	110000	6 ABA92787_1	Continuation (2 of
39	843	38.3	1080	13 ADT46404	Adt46404 Bacterial
40	840.5	38.2	1071	5 AAF94408	Aaf94408 Haemophil
41	840.5	38.2	1071	6 ABK64972	Abk64972 DNA encod
42	840.5	38.2	1074	4 AAS53247	Aas53247 Haemophil
43	840.5	38.2	1074	8 ACA34045	ACA34045 Prokaryot
44	840.5	38.2	1077	13 ADT45862	Adt45862 Bacterial
45	840.5	38.2	110000	2 AAT42063_02	Continuation (3 of

## ALIGNMENTS

RESULT 1  
AAZ50438  
ID AAZ50438 standard; cDNA; 1626 BP.

AC AAZ50438;

DT 18-MAY-2000 (first entry)

DE Corn clone csiln.pk0050.d11 encoding chorismate synthase.

KW Corn cDNA clone csiln.pk0050.d11; chorismate synthase;  
immunological screening; herbicide resistance; antibody; gene mapping;  
corn; ss.

OS Zea mays.

FH Key Location/Qualifiers  
CDS 79..1368

FT /\*tag= a

FT /product= "Chorismate synthase"

PN WO200005353-A2.

XX 03-FEB-2000.

PF 20-JUL-1999; 99WO-US016353.

XX 21-JUL-1998; 98US-0093611P.





KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 200SEP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0128845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147152P.  
PR 05-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147433P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.



```
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 9,98e-125 Length: 1595
Score: 1461.00 Matches: 276
Percent Similarity: 81.91% Conservative: 59
Best Local Similarity: 67.48% Mismatches: 62
Query Match: 66.41% Indels: 12
DB: 3 Gaps: 2

US-10-677-179-8 (1-429) x AAC40227 (1-1595)
QY 30 ProAlaSerSerSerLeuArgPheAlaValHisArg----- 41
DB 176 CCGGGTCTCTCTCTCTCCCTCGGAGCTCCGTCGTCCTTCTCTCCCGCGCTTCAGATC 235
QY 42 ---CysArgThr-----AlaArgLeuGluValLysAlaSerGlyAsnThrPheGly 57
DB 236 TCTCTCCGTACCAACAGGAGAACTTCCAGATACAAAGCTACTGGAAGTTCATATGGG 295
QY 58 AsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIle 77
DB 296 ACTCATTTTCGAGTTTCAACTTTTCGAGAATCACATGGAGGAGGAGTTGTTGTATCAT 355
QY 78 SerGlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArg 97
DB 356 GATGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 415
QY 98 ArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIle 117
```

```
DB 416 AGGAGACCTGGTCAGAGCAGGATCAACTCTCTAGAAAGGAGACTGATCTTCCGGATA 475
QY 118 LeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsn 137
DB 476 TCTTCTGGAGTCTCGAAAGGAATGACGACGAGCAACACCTATCCATGTGTGTGTGCTAAC 535
QY 138 ThrAspGlnIleGlySerAspHisArgGluIleAlaAenValTyrArgProSerHisAla 157
DB 536 ACAGATCAGAGGACTTGAATACAGTGAATGTGGTTCCTATAGACCATCGCATGCT 595
QY 158 AspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyGlyArgSerSer 177
DB 596 GATGCAACTTATGACATGAAGTATGTGTGCAGATCAGTGCAGGCTGGAGGAAGATCTTCA 655
QY 178 GlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeu 197
DB 656 GCTAGAGAGACCATTGGAGAGATTGCTCTCTGGAGCTTTGGCCAGAAAAATTTTGAAGCAA 715
QY 198 LysCysGlyLeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGlu 217
DB 716 TTTGAGGAACCTGAGATTCTTGCTATGTCTCGCAAGTTCACCATGTTGTACTTCCAGAA 775
QY 218 AspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCys 237
DB 776 GAATTTGGTAGACACGAGCAATTTAACTCTCGAAGATAGAAAAATAACATGTCAGATGC 835
QY 238 ProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGly 257
DB 836 CCTAATCCGAGTATGCGAAAAAGATGATAGTGCATGATGATGCTGTCTCAGGACAAAAGG 895
QY 258 AspSerValGlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySer 277
DB 896 AACTCTGTGTGGTGGTGTGTGACCTGCATTTGCGGAATGCTCCACGTGGGGCTTGGTACA 955
QY 278 ProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSer 297
DB 956 CCGGTTTTCGATAAATTTGAAGCAGAACTGGCAAAAGCTTGTATGTCGTACTTCCACACA 1015
QY 298 AsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsn 317
DB 1016 AAGGGATTTGAGTTTGGAAAGCGGCTTTGCAGGTACCTTTTTCACCTGGTCTTGAACACA 1075
QY 318 AspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGly 337
DB 1076 GATGAGTTCTATACCGATGAATAATGGAAGATACGTACCAAGCAAAACCGATCTGGTGA 1135
QY 338 ValGlnGlyGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThr 357
DB 1136 ATTCAGGAGGAGATCTCAATGGTGAATATAAATACATGAGAGTAGCTTCAAGCCACACA 1195
QY 358 ProSerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeu 377
DB 1196 TCAACAATTCGAAGGAGCAGATACGCTAACCCAGAGACAGAGTGAAGACCGAATGAT 1255
QY 378 AlaArgGlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMet 397
DB 1256 GCGGTGTGTCTCATGATCTTGTGTGTTCTCTCGAGTGTGCCAATGGTGAATCAATG 1315
QY 398 AlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAla 417
DB 1316 GTGGCTTTAGTCTTGTGTGATCAATTTGATGGCGCAATACGACCAATGTCATTTGTTTCCA 1375
QY 418 LeuAsnThrAlaLeuGlnGluProVal 426
DB 1376 ATAAATCCAGAGTTGCAGGAACCTCTC 1402
RESULT 4
ADN73862
ID ADN73862 standard; cDNA; 1311 BP.
XX
AC ADN73862;
XX
```





```

AAZ50440
ID AAZ50440 standard; cDNA; 966 BP.
XX
AC AAZ50440;
XX
DT 18-MAY-2000 (first entry)
XX
DE Rice clone rls72.pk0029.g8 encoding chorismate synthase.
XX
KW Rice cDNA clone rls72.pk0029.g8; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW rice; ss.
XX
OS Oryza sativa.
XX
FH Key
FT CDS 45..964
FT /tag= a
FT /product= "Chorismate synthase"
FT /note= "no stop codon given"
XX
PN WO200005353-A2.
XX
XX 03-FEB-2000.
XX
XX 20-JUL-1999; 99WO-US016353.
XX
XX 21-JUL-1998; 98US-0093611P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon RE, Falco SC, Pember SO;
XX
XX WPI; 2000-182687/16.
XX
XX P-PSDB; AAY44895.
XX
XX New chorismate synthase polypeptides used to alter the level of the
XX enzyme and thus the level of aromatic to non aromatic amino acids in
XX transformed plants.
XX
XX Claim 3; Page 35; 39pp; English.
XX
XX The present sequence is rice cDNA clone designated rls72.pk0029.g8
XX encoding chotimamate synthase. This rice is isolated from rls72 cDNA
XX library which is derived from rice leaf 15 days after germination, 72
XX hours after infection of strain Magaportha grisea 4360-R-67. Chimeric
XX gene comprising this cDNA operably linked to regulatory sequences is used
XX to transform host cells to alter the level of expression of chorismate
XX synthase. The gene and its products may be used for immunological
XX screening of cDNA expression libraries and to create transgenic plants
XX which may also be herbicide resistant. Synthetic peptides derived from
XX the gene are to raise antibodies, and used in screening assays to
XX identify inhibitors which may be useful as herbicides. Probes and primers
XX from chorismate synthase gene are used for gene mapping and as markers
XX for traits linked to those genes
XX
SQ Sequence 966 BP; 251 A; 234 C; 264 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,26e-106 Length: 966
Score: 1257.00 Matches: 245
Percent Similarity: 89.30% Conservative: 22
Best Local Similarity: 81.94% Mismatches: 26
Query Match: 57.14% Indels: 6
DB: 3 Gaps: 2

US-10-677-179-8 (1-429) x AAZ50440 (1-966)
QY 15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro 30
Db 75 GCGCGGTCCTCCCGCGCGCGCGCGCGGTTCCGCGGCTTCCCGAGTCCGCGCG 134
QY 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuValIys 50

```







Qy 413 CysgluWet 415  
Db 1075 TGTAAAGGTG 1083

RESULT 9  
ADT45459  
ID ADT45459 standard; cDNA; 1086 BP.  
XX  
AC  
AC ADT45459;  
XX  
XX  
02-DEC-2004 (first entry)  
XX  
XX  
DE Bacterial polynucleotide #20210.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polynucleotide; gene; ss.  
XX  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
XX  
PD 18-DEC-2003.  
XX  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
XX  
PR 21-FEB-2002; 2002US-0360309P.  
XX  
XX (CAOX/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX  
PS Claim 1; SEQ ID NO 43997; 122pp; English.  
XX  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polynucleotide used in  
CC the scope of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX



```
Db 895 GCATTATAGCCGACAGCAACATTAGAAAAGAACAGAGACTGTACTCGTGAGGCCGAA 954
Qy 373 AenValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
Db 955 GAAACACATTATTAGCAGCAAAAGGACGACACGATCCTTGTTGTTATTACCGCGTCAGTGCCA 1014
Qy 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db 1015 ATGGTTGAGGCAATGGTCGCTGGTACTGTATGTGACCATTTGTTACGGCATCATGGTCAG 1074
Qy 413 CysGluMet 415
Db 1075 TGCAAAGTC 1083

RESULT 10
AAZ50437
ID AAZ50437 standard; cDNA; 1015 BP.
XX
AC AAZ50437;
XX
DT 18-MAY-2000 (first entry)
XX
DE Wheat clone wreln.pk0094.e6 encoding chorismate synthase.
XX
KW Wheat cDNA clone wreln.pk0094.e6; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW wheat; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 2..775
FT FT /*tag= a
FT TT /product= "Chorismate synthase"
XX
PN WO200005353-A2.
XX
XX
XX 03-FEB-2000.
XX
XX 20-JUL-1999; 99WO-US016353.
XX
XX 21-JUL-1998; 98US-0093611P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon RE, Falco SC, Pember SO;
XX
XX WPI; 2000-182687/16.
XX
XX P-PSDB; AAY4892.
XX
XX New chorismate synthase polypeptides used to alter the level of the
XX enzyme and thus the level of aromatic to non aromatic amino acids in
XX transformed plants.
XX
XX Claim 3; Page 31; 39pp; English.
XX
XX The present sequence is wheat cDNA clone designated wreln.pk0094.e6
XX encoding chorismate synthase. This clone is isolated from wreln cDNA
XX library which is derived from wheat root from 7 day old etiolated
XX seedling. Chimeric gene comprising this cDNA operably linked to
XX regulatory sequences is used to transform host cells to alter the level
XX of expression of chorismate synthase. The gene and its products may be
XX used for immunological screening of cDNA expression libraries and to
XX create transgenic plants which may also be herbicide resistant. Synthetic
XX peptides derived from the gene are to raise antibodies, and used in
XX screening assays to identify inhibitors which may be useful as
XX herbicides. Probes and primers from chorismate synthase gene are used for
XX gene mapping and as markers for traits linked to those genes
XX
XX Sequence 1015 BP; 283 A; 213 C; 251 G; 268 T; 0 U; 0 Other;
SQ
Alignment Scores:
```

```
Pred. No.: 1.39e-88 Length: 1015
Score: 1066.00 Matches: 201
Percent Similarity: 93.83% Conservative: 27
Best Local Similarity: 82.72% Mismatches: 15
Query Match: 48.45% Indels: 0
DB: 3 Gaps: 0

US-10-677-179-8 (1-429) x AAZ50437 (1-1015)
Qy 186 AlaAlaGlyAlaLeuProLysLysIleLeuLysLysCysGlyLeuGluIleLeuSer 205
Db 8 GCTCAGGAGCTGTTGCAAGAAATTTCTTAAGCTGAATGTGGAGTAGAGATTCTAGCA 67
Qy 206 PheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerVal 225
Db 68 TTTGTTTCCAAAGTGCATCAAGTGTACTTCTCTGACACGCGATTGATTATGAACCTCTT 127
Qy 226 ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys 245
Db 128 ACCCTGGATCAGATAGAGAGCAACATTTGTAGATGTCTGATCCAGAAATATGCACAGAAG 187
Qy 246 MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThr 265
Db 188 ATGATTGATGCAATTGATAAAGTACGAGTTAATGGGAATTCGATTGGTGGGGTGGTCACA 247
Qy 266 CysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSer 285
Db 248 TGCATTGCCAGAAATGTTCTCTCGTGGGCTGGCTCTCTGTATTGTGACAACTTGAAGCT 307
Qy 286 GluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGly 305
Db 308 CTACTGGCAAGGCTATGCTTTCTTCTTCTGCAAGCAAGGGGTTTGAGATCGGTAGTGA 367
Qy 306 PheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAla 325
Db 368 TTTGACGGTACTGACCTAACTGGAAGTGAGCATTAACGATGATGTTCTATATGACGAGGCT 427
Qy 326 GlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnVal 345
Db 428 GGAATGTAAAGACACAGAACCAATCGCTCGGGCGGTGTACAGGGAGGGATATCAATGGT 487
Qy 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365
Db 488 GAAACTATATACTTCAAAAGTAGCTTTCAAGCCACACAGCAACTATTGGGAAGAGCAAAAT 547
Qy 366 ThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCys 385
Db 548 ACTGTAAACAGGGGATCATGAGGATATCGAACTTCTGCAAGGGGTGCGCCATGACCCATGT 607
Qy 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405
Db 608 GTGTCCTCTCGGGCTGTTCGAATGGTGGAGACGATGGCTGCTATGGTCTCTATGGACACAG 667
Qy 406 LeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAenThrAlaLeuGlnGluPro 425
Db 668 CTGATGGCACATGTTGCTCAGTGGCAGATGTTCCGCGCTGAACCTCGCCCTCAAGAACA 727
Qy 426 ValGlySer 428
Db 728 ATCGGCTCC 736

RESULT 11
ADT46302
ID ADT46302 standard; cDNA; 1080 BP.
XX
AC ADT46302;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #21053.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
```

pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbon; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.

**Bacte:**

US2003233675-A1.

18-DEC-2003

20-FEB-2003: 2003UIS-00369493

21-FEB-2003: 2003US-0360030B

(C) 2000 by CRC Press LLC

(CAOY/) CAO Y.  
(HINX/) HINX E C Y

(HINK/) HINKLE G J.  
(STAT/) STATEP S C

(SLAT/) SLATER S C.  
(CHEN/) CHEN Y

(CHEN/) CHEN X.  
(GOLD/) GOLDMAN B S

Gao Y Hinkle GT Slater SC Chen Y Coldman PG.

WPT - 2004-2006

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

Claim 1: SEQ ID NO 44740: 122bp: English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from [uspto.gov/sequence.html](http://uspto.gov/sequence.html).

Sequence 1090 BP: 204 A: 331 C: 369 G: 176 T: 0 U: 0 Other:

**Alignment Scores:**

Pred. No.:	1.09e-85	Length:	1080
Score:	1035.00	Matches:	600
Percent Similarity:	72.65%	Conservative:	23
Best Local Similarity:	55.25%	Mismatches:	95
Query Match:	47.05%	Indels:	4
DB:	13	Gaps:	2

U.S.-10-677-179-8 (1-429) x ADT46302 (1-1080)

Nv 53 G] vAgnThrPheG] vAgnThrPheG] nVg] x] eThxThyrc] m] "SgrH; eG] m] m] y 72

[illegible]

QY	73	ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln	92
DB	64	GTGGCGTGTATGATGGAGGGCTGCGCTTCCACGCGCTGATCTGGATGTGTGACGCATCCAG	123
QY	93	ValGluLeuLeuArgArgArgProGlyGlnSerArgIleThrSerThrArgIleGluThr	112
DB	124	GCGGAGTTGGATGGCGACCGCGGGTTCAGAGCCACATCACACCCACCCACCAAGAACGCC	183
QY	113	AspThrCysIleLeuSerGlyThrHisGluGlyValThrThr---GlyThrProIle	131
DB	184	GATCAGTGGAGGTCTCAGGGCCCTTCGTATGGGCGAGACCACTCTGGGAAACACCGATC	243
QY	132	LeuValIleValProLeuThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal	151
DB	244	GCCATGTTGTGGCGCAACAAGGACGAGCGCCCGGCGACTACAAAGGACATGGCGGTGGCG	303
QY	152	TyrArgProSerHisAlaAspAlaThrTyrAspPheIleGlyValArgAlaValGln	171
DB	304	TTTCGGCCCTTCTCATGCCGATGCCACTCACCGGTGAATAACCGCGTGCAGGCCCGCAGT	363
QY	172	GlyGlyGlyArgSerSerGlyArgIleGlyThrValGlyArgValAlaGlyAlaLeuPro	191
DB	364	GTTGGCGGGGGGCTCAGCCCGGAAACAATTTGGGCGTGGCGCGCGGTGCATATCGCC	423
QY	192	LysIleLeuIleLysLeuIleCysGlyLeuGluIleLeuSerPheValSerIleValHis	211
DB	424	AAGCAACTGTGGCAAGCGCGCAGGACGAGAGTCTTGGCCCTGGGTCAAACGCATCCAC	483
QY	212	GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluIleGlu	231
DB	484	ACCATC-----GAGGCGAATGTGGATCCCGAGCGGTAACTGGTGAAGCGCATCGAG	534
QY	232	SerAsnIleValArgCysProAspProGluTyrAlaGluIleMetIleAspAlaIleAsp	251
DB	535	AGCAACATCTTTCGTCCCCGGATGCCCTCACAGCCGCTCAGATGTGTGGAAACGGATCGAG	594
QY	252	ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnVal	271
DB	595	GCGATCGGGCGCAGCGCGACTCATGTGTGGCGGTGATCGAATCGTGGTGGTAAACCT	654
QY	272	ProArgGlyLeuGlySerProValPheAspIleLeuGluSerGluLeuAlaIleValMet	291
DB	655	GCTCCGGGGCTGGGTATGCCCTGTGTTCGACAAAGCTGGAGCCGATCTGGCCAAAGCGGTG	714
QY	292	LeuSerIleProAlaSerAsnGlyPheGluIleGlySerClyPheAlaGlyThrAspLeu	311
DB	715	ATGTCCTCCCCGCCCAACCAAGGCGTTTGAGATCGGTCTGGCTTCAGCGGAAACCCCTGCTC	774
QY	312	ThrGlySerGluHisAsnAspGluPheTyrMetAspIleAlaGlySerValArgThrArg	331
DB	775	AAGGGCAGCGAGCACACGATGCTTTGTGCCCCACCGATGATCGCGGTGCAGACGGCC	834
QY	332	ThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLys	351
DB	835	ACCAACAACCTCAGTGGCATCCAGGCGCGCATCAGCAACCGCGAGCCGATTGTGATCCGG	894
QY	352	ValAlaPheIleYsProThrProSerIleGlyValIleGlnAsnThrValSerArgGluArg	371
DB	895	GTGGCGTTCAGGCCCAACCGCCAGATCCCGCAAGGAACAGCAGACCAATTGATCCGACGGC	954
QY	372	GlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal	391
DB	955	AAGCCACCACTTGGCGCGCAGGGTGCACACGACCCCTGTGTGTGTGTCGAGGGCGGTG	1014
QY	392	ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla	411
DB	1015	CCGATGTGTGAGGCGATGTGGCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1074
QY	412	GlnCys	413
DB	1075	CAGTGC	1080

## RESULT 12

AD58432  
ID AD58432 standard; cDNA; 1068 BP.  
AC AD58432;  
XX

02-DEC-2004 (first entry)

Bacterial polynucleotide #10419.

Recombinant DNA construct; transformed plant; improved plant property;  
cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
pathogen tolerance; pest tolerance; plant disease resistance;  
cell cycle pathway modification; plant growth regulator;  
homologous recombination; seed oil yield; protein yield; carbohydrate;  
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
bacterial polynucleotide; gene; ss.

Bacteria.

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

(CAO/Y) CAO Y.  
(HINK/) HINKLE G J.  
(SLAT/) SLATER S C.  
(CHEN/) CHEN X.  
(GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide  
for expression of a polynucleotide encoding a polypeptide from a  
microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 34106; 122pp; English.

The invention relates to a recombinant DNA construct comprising a  
promoter functional in a plant cell, where the promoter is positioned to  
provide for expression of a polynucleotide encoding a polypeptide from a  
microbial source. The invention also relates to a transformed plant  
comprising the recombinant DNA construct and a method of producing a  
transformed plant having an improved property. The plant is a crop plant  
such as maize or soybean. The method of producing a transformed plant  
having an improved property comprises transforming a plant with the  
recombinant DNA construct and growing the transformed plant, where the  
polynucleotide or polypeptide is useful for improving plant properties.  
The recombinant DNA construct is useful for producing plants with  
improved plant properties, e.g. improved cold, heat or drought tolerance,  
tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
increased resistance to plant disease, better growth rate by modification  
of the cell cycle pathway with plant growth regulators, increased rate of  
homologous recombination, modified seed oil or protein yield and/or  
content, improved yield by modification of carbohydrate, nitrogen or  
phosphorus use and/or uptake, by modification of photosynthesis or by  
providing improved plant growth and development under at least one stress  
condition, improved lignin production or improved galactomannan  
production. This sequence represents a bacterial polynucleotide used in  
the scope of the invention. Note: The sequence data for this patent did  
not form part of the printed specification but was obtained in electronic  
format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

Sequence 1068 BP; 325 A; 209 C; 266 G; 268 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.02e-80 Length: 1068  
Score: 981.00 Matches: 187

Percent Similarity:	71.51%	Conservative:	64
Best Local Similarity:	53.28%	Mismatches:	96
Query Match:	44.59%	Indels:	4
DB:	13	Gaps:	2
US-10-677-179-8 (1-429) x ADS58432 (1-1068)			
QY	54	AsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyVal	73
DB	7	AATACATACGCACAAATATTTCGCATTACAACTTTGGAGATCATCATGGTAATGCTGTT	66
QY	74	GlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAapLeuGlnVal	93
DB	67	GGAGTTATAGTGGATGTTGTCGCCCTAATATCGAATCGACGAAGCGTTTATTCAATCT	126
QY	94	GluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp	113
DB	127	GAATGGCCGCGCAGAAACCCGGTCAGTCCAAAGATCGTTACACAGCGTCAGGAAGCGAT	186
QY	114	ThrCysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeuVal	133
DB	187	GAGATAGAAATTTATCCGGTGATTTGAGGAAATCTACCGGTACGCTATTGCTATG	246
QY	134	IleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArg	153
DB	247	ATGGTGGCGTAACGCGATCAGCGCAGTAAAGATTATTACACATTGCAGATCGTTCCGC	306
QY	154	ProSerHisAlaAspAlaThrTyrAspPheIleThrGlyValArgAlaValGlnGlyGly	173
DB	307	CCGTCCTCATGCAGATTACACCTATCAGGAAATAATATGCTTCCTCGATTATCCGGCGGC	366
QY	174	GlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyValAlaLeuProLysLys	193
DB	367	GGCAGAAAGTTCTGCCGGGAAACCGCTGCACCGCTAGCTGCTGGCGCACTGGCTAAATG	426
QY	194	IleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal	213
DB	427	ATTCTG---GCTAAAGTAAATATTAAAGATTTCAGGCATATGTTTCACAGTTGGTCCGTTG	483
QY	214	ValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSerAsn	233
DB	484	AAACTT-----ACGACCGATTATAAAGCATGTAATATTCCGAAACCGGAAACCAAT	534
QY	234	IleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArgVal	253
DB	535	ATCATCCTGTTCCGATGATGCTGTGACGACCAAAATGATTTGATTTATGATGAACA	594
QY	254	ArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnValProArg	273
DB	595	CGTAAATAAAGATATACCATTTGGTGGTGTGTAAGCTGCGTAATTTGAAGGTGTTCCGGTT	654
QY	274	GlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSer	293
DB	655	GGTTTAGCGCAACCGGTATTTGATAAACTGCATCAGAACTTGGCAAAACCATGTTAAGC	714
QY	294	IleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGly	313
DB	715	ATCAATGCCGTTAAAGGGTTTGAATACGGAAGCGGTTTGAAGGTGTAGACAGTTCCGG	774
QY	314	SerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsn	333
DB	775	TCACAGCAACAGTAAATTTTATGCGGATGAACAGGGCAATGTACGTACACAAACCAAC	834
QY	334	ArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPheLysValAla	353
DB	835	AATTACGCGGTATTTCAGGCGGTATCAGCAACGGAGAGATATTTATTTCGCGGTGCA	894
QY	354	PheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsn	373
DB	895	TTTAAACCTGTAGCTACATTTATGTTGACCAAGAAAGCATCAACCAAGTGAACG	954
QY	374	ValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaValProVal	393

[illegible]

CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX  
SQ  
Sequence 1131 BP; 312 A; 267 C; 272 G; 280 T; 0 U; 0 Other;

Alignment Scores:		
Pred. No.:	1,01e-76	Length: 1131
Score:	338.00	Matches: 183
Percent Similarity:	68.12%	Conservative: 67
Best Local Similarity:	49.86%	Mismatches: 109
Query Match:	42.64%	Indels: 3
DB:	13	Gaps: 8

US-10-677-179-8 (1-429) x ADT47171 (1-1131)

[illegible]



```
QY 74 GlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGlnVal 93
DB 64 GGGTGTATTGTTGATGGTGTCCACAGGATGTCATTAACTGAGCAGATATTCACCA 123
QY 94 GluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp 113
DB 124 CAATTGACTCGTAGAGTCCAGGTCAAAGTAAATTTATCTACCCCAAGAGATGAAAGAT 183
QY 114 ThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeuVal 133
DB 184 TTAGTTGAATAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 243
QY 134 IleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArg 153
DB 244 ATTGTGAGAAATAAAGATCATCTCTGCTGATGATGATGATGATGATGATGATGATG 303
QY 154 ProSerHisAlaAspAlaThrThrAspPheLysTyrGlyValArgAlaValGlnGlyGly 173
DB 304 CCAAGTCATGCCGATTCGACCTATATACAGAAATACGGAACCAAAATCTAGTTCTGCTG 363
QY 174 GlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaAlaLeuProLysLys 193
DB 364 GGTAGATCATCTCTGAGAGAACTATTTGGTAGAGTTGCCGACGAGCCATTGCGAGAAAG 423
QY 194 IleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal 213
DB 424 ATTTTGGCCAAAGTGAATAATGTTGAATGTTGCTTTGTTAGTCTATTTGTTGAATA 483
QY 214 ValLeu-----ProGluAspAlaValAsp-----TyrGlySerValThrLeu 227
DB 484 TCTATGAGTAAATCTCTCAAGATGCAAAATTCAAAGAACTTTTAAACACTACTACCTAGA 543
QY 228 GluGlnIleGluSer---AsnIleValArgCysProAspProGluTyrAlaGluLysMet 246
DB 544 GAACAAGTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 603
QY 247 IleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCys 266
DB 604 GTTAAAGTGTGAAATAATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 663
QY 267 ValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu 286
DB 664 GTTATCAGAAATGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATG 723
QY 287 LeuAlaIleAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPhe 306
DB 724 TTGGCTCATGCCATGTTATCATATCCCGCTACCAAGGGGTTTGAATTTGGTTCTGATTT 783
QY 307 AlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLys---Ala 325
DB 784 GAAGGTATCAAGATCCCTGTTCAAAACATACAGATGATGATGATGATGATGATGATGAT 843
QY 326 GlySerValArgThrArgThrAsnArgSerGlyValGlnGlyIleSerAsnVal 345
DB 844 GGAAGATTAAAGACTGAACAAACAATAGTGTGTTGATGATGATGATGATGATGATGATG 903
QY 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365
DB 904 GAAACATTTATTTCTCAGTTGCTTCAATTCGATGCTGATGATGATGATGATGATGATG 963
QY 366 ThrValSerArgGluArgGlnAsnValGluLeuAlaArgGlyValArgHisAspProCys 385
DB 964 ACTGCCCATATGATGTTAAAGTGGTGTGTTTGGCTGATGATGATGATGATGATGATGAT 1023
QY 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405
DB 1024 GTCAACCCCAAGACAGCTTCAATTTGCGAAGCCATGATGATGATGATGATGATGATGAT 1083
QY 406 LeuMetAlaHisValAlaGln 412
DB 1084 TATATGATTCACCAAGCTAGA 1104
```

## RESULT 15

```
ADR85489
ID ADR85489 standard; DNA; 1239 BP.
XX
AC ADR85489;
XX
DT 04-NOV-2004 (first entry)
XX
DE Aspergillus fumigatus essential gene open reading frame #126.
XX
KW Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;
XX drug screening; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO2004067709-A2.
XX
PD 12-AUG-2004.
XX
PF 16-JAN-2004; 2004WO-US001099.
XX
PR 17-JAN-2003; 2003US-0441281P.
XX
PR 13-JUN-2003; 2003US-0478196P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PA (ELIT-) ELITRA CANADA LTD.
XX
PI Jiang B, Hu W, Lemieux S, Roemer T;
XX
DR WPI; 2004-S94200/57.
XX
DR P-PSDB; ADR86076.
XX
PT New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT encoding a gene product, useful for diagnosing and/or treating invasive
PT fungal infections, such as Farmer's lung disease.
XX
PS Claim 2; SEQ ID NO 2126; 164pp; English.
XX
CC The present invention relates to Aspergillus fumigatus genes that are
CC essential and are potential targets for drug screening. The methods and
CC compositions of the present invention are useful for diagnosing and/or
CC treating invasive Aspergillus fumigatus infection, including the allergic
CC forms of the disease, such as Farmer's lung disease. They can also be
CC used in various drug discovery purposes, such as expression of the
CC recombinant protein, hybridization assay and construction of nucleic acid
CC arrays. The present sequence represents an Aspergillus fumigatus
CC essential gene open reading frame, used during diagnosis and drug
CC development in the invention. These genes share a high degree of sequence
CC conservation with known essential genes of candida albicans. The sequence
CC data for this patent is not represented in the printed specification, but
CC was obtained in electronic format from WIPO.
XX
SQ Sequence 1239 BP; 255 A; 431 C; 330 G; 223 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6 9e-76 Length: 1239
Score: 929.50 Matches: 196
Percent Similarity: 66.49% Conservatives: 58
Best Local Similarity: 51.31% Mismatches: 105
Query Match: 42.25% Indels: 23
DB: 13 Gaps: 6
US-10-677-179-8 (1-429) x ADR85489 (1-1239)
QY 54 AsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyVal 73
DB 4 TCGACGTGGGAGAGATCTTTTCGGGTCCACCATGATGTTGAATCCATTCGCGCTCTGTC 63
QY 74 GlyCysValIleSerGlyCysProArgIleThrProLeuThrGluAlaAspLeuGlnVal 93
DB 64 GGTCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 94 GluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp 113
```



Job time : 639 secs

```
Db      124  CAGATGACTCGAAGACGTCCTCCGGGAGAGTGTCTTAACGACGCTCGAAATGAAAAAGAC 183
QY      114  ThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeuVal 133
Db      184  CGAGTAGAGATCCAGTCTGGAACGAGGTTCGGCATCACCTGGGTACCCCGATTGGAATG 243
QY      134  IleValProAsnThrAspGlnIleGlySerAspHis---ArgGluIleAlaAsnValTyr 152
Db      244  ATGGTGGCGCAACGAGGATCAGAGACCCAGGACTACGGTGGCGAGCACCAATGGATCTCTAC 303
QY      153  ---ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
Db      304  CCTCGTCCCGAGTCAGCTGATTACTTACCTGGAGAAATACGGTGTCAAGCGGAGCAGC 363
QY      172  GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
Db      364  GGTGGTGGCCGGAGTAGTCCCGCGAGACCATTTGGCGGTTCGCCGAGGAGCCATTGGC 423
QY      192  LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db      424  GAGAAAGTACCTACGGCTGTCGATGGTGTGGAATTTGTGGCCTTTGTCTCCGTTGGT 483
QY      212  GlnValValLeuProGluAspAlaValAspTyrGlySerValThr----- 226
Db      484  AAGAACACCTTTTCCCGCGAGCCCCCGAGACCCCTTCTCCATGACCAACCCCTGAGTTC 543
QY      227  -----LeuGluGlnIleGluSerAsnIleVal-----ArgCys 237
Db      544  CTGAAGCTCATCGAGACCATCGACCTAAGACTGTCGATGCTTCGTCCCCACTCGCTGC 603
QY      238  ProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGly 257
Db      604  CCGAACGAGGAGCGCGCGACGACATGACAAAGTGTGATCGAGACTTTTCGGGACCAACAA 663
QY      258  AspSerValGlyGlyValIleThrCysValAlaIleArgAsnValProArgGlyLeuGlySer 277
Db      664  GATAGCATCGCGGACCGCTACCTCGGTGATCCGCAAGCTCCCGTGGCCCTGGGCGAG 723
QY      278  ProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSer 297
Db      724  CCTTGCTTCGACAAGCTCGAGGCCAAGCTGGCGCACGCCCATGCTCAGCATCCCGCCACC 783
QY      298  AsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsn 317
Db      784  AAGCGCTTTGAGATCGGCTCGGGCTTCGGTGGCTCGAGGTCCCGGGCTTCATCCACCAAC 843
QY      318  AspGluPheTyrMetAsp-----LysAlaGlySerValArg-----Thr 330
Db      844  GACCTCTTCCCGTCTCCGAGGTCCAGACCCGCGGCGAGCACAGCGGCTTGACCACC 903
QY      331  ArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPhe 350
Db      904  AAGACCAACAACACTCCGGCGGCATCCAGGGCGGATCTCCACGGCGCTCCCATCTATTTC 963
QY      351  LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db      964  CGCGTTGCCCTTCAAGCCCCCGCCACCATCGGCCAAGGCTCAGACCAACCGCCTCTTACAGC 1023
QY      371  ArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db      1024  TTCGAGGAGGGGATCTCCGAGGCCAAGGCCCGCCAGACCCCTGGTTACCCCTCGTGCT 1083
QY      391  ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db      1084  GTCCCCATCGTCGAGGCCATGTCGGCCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1143
QY      411  AlaGln 412
Db      1144  GCCCGC 1149
```

Search completed: August 27, 2005, 04:55:06



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 27, 2005, 04:20:24 ; Search time 3357 Seconds  
(without alignments)

864.335 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10677179/runat\_25082005\_150240\_24847/app\_query.fasta\_1.583  
-DB=EST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10677179 @Cn1\_1\_5180 @runat\_25082005\_150240\_24847 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

```

1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	3 AY105171	AY105171 Zea mays
2	1441	65.5	1488	3 CINSOACWT	BX815746 Arabidops
3	1434	65.2	1494	3 CINSOACWM	BX818532 Arabidops
4	1398	63.5	1538	3 CINSOACVT	BX814847 Arabidops
5	1182.5	53.8	1160	7 CK210372	CK210372 FGAS02217
6	1170	53.2	923	7 CO082566	CO082566 GR_Ea47E
7	1135	51.6	763	6 CB654761	CB654761 OSJNEC071
8	1120.5	50.9	911	7 CV256756	CV256756 WS0244.B2
9	1087	49.4	865	7 CO114396	CO114396 GR_Eb015

10	1042	47.4	786	4 BG600177	BG600177 EST505072
11	1039	47.2	835	4 BM411032	BM411032 EST585359
12	1032	46.9	710	4 BJ268020	BJ268020 BJ268020
13	1027.5	46.7	814	7 CN144654	CN144654 WOUND1_23
14	1022	46.5	723	7 CF303160	CF303160 ABF1--01-
15	1022	46.5	755	7 CK279831	CK279831 EST725909
16	1018	46.3	835	7 CO080629	CO080629 GR_Ea44D
17	1011	46.0	702	5 BU099249	BU099249 WHE3305.C
18	996.5	45.3	919	7 CN385253	CN385253 LE2TR03E1
19	990	45.0	869	7 CO093258	CO093258 GR_Ea14L
20	989	45.0	864	6 CA189902	CA189902 SCCCLR1C0
21	988	44.9	803	6 CB066694	CB066694 EST645317
22	985	44.8	690	6 CA124299	CA124299 SCQGLR108
23	984	44.7	667	6 CA119896	CA119896 SCCCLR106
24	970	44.1	818	7 CO132399	CO132399 GR_Eb45D
25	967	44.0	694	6 CD870292	CD870292 AZO2_113P
26	959	43.6	812	7 CN129049	CN129049 RHOH1_33
27	954	43.4	775	7 CO101191	CO101191 GR_EB002
28	948	43.1	696	5 BQ841547	BQ841547 WHE4212.D
29	946	43.0	783	7 CO112743	CO112743 GR_Eb004
30	932	42.4	871	7 CO090892	CO090892 GR_Ea11B
31	929	42.2	673	7 CF571696	CF571696 MCS016H05
32	924	42.0	689	6 CA298301	CA298301 SCRFPL803
33	921	41.9	726	6 CD008025	CD008025 VVB082A11
34	918	41.7	639	4 BG414940	BG414940 HVSMEX000
35	918	41.7	722	4 BI930866	BI930866 EST550755
36	916.5	41.7	749	6 CA094750	CA094750 SCCCL401
37	916	41.6	957	7 CV263720	CV263720 WS02022.B
38	908	41.3	712	6 CD712988	CD712988 VVB213A07
39	906	41.2	708	6 CD009988	CD009988 VVB102E11
40	901	41.0	682	2 AW223745	AW223745 EST300556
41	901	41.0	707	6 CD714440	CD714440 VVB200H12
42	895	40.7	605	6 CA220612	CA220612 SCRFPL400
43	893	40.6	907	7 CO074131	CO074131 GR_Ea34B
44	892	40.5	661	6 CD899216	CD899216 G174_111J
45	891	40.5	629	1 AV832717	AV832717 AV832717

## ALIGNMENTS

RESULT 1	AY105171	AY105171	1626 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	Zea mays	PCO087106	mRNA sequence.			
DEFINITION	AY105171					
ACCESSION	AY105171					
VERSION	AY105171.1	GI:21208249				
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE						
AUTHORS	Hayney,C.P., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2 (bases 1 to 1626)					
AUTHORS	Coe,E.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.					

FEATURES  
source

Location/Qualifiers  
1..1626  
/organism="Zea mays"

```

/mol_type="mrna"
/db_xref="MaizeDB:634903"
/db_xref="caxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 3,826-225 Length: 1626
Score: 2200.00 Matches: 429
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-677-179-8 (1-429) x AY105171 (1-1626)
QY 1 MetThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
DB 79 ATGACGCGTGTCCCAAGCCACAGCAGGTGGGCACTCAGGGCAGGCTCGACCCCGC 138
QY 21 AlalLeGlyAlaLeuLeuGluPheAlaProLysSerSerLeuArgPheAlaValHis 40
DB 139 GCGATCGGCGCTTCTGCGAGTTTGGCCCGCAGCTCTCTCCCTCGGCTTGGCGGTGCAC 198
QY 41 ArgCysArgThrAlaArgLeuGlnValLysAlaSerGlyAsnThrPheGlyAsnThrPhe 60
DB 199 CGCTCGCGCAGCTGCTGGCTAGAGTGAAGGCATCTGGAACACAGTTTGGAAACTACTTT 258
QY 61 GlnValAlaThrTyrGlySerHisGlyGlyValGlyCysValIleSerGlyCys 80
DB 259 CAGGTTCACACCTATGGTGAATCTCATCGGGGTGGTGGTGGTGGTATCAGTGGTGT 318
QY 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
DB 319 CCACCTAGAAATTCACCTCACTGAGGACAGCCTTACAAAGTTGAACCTGATCGAAGACGGCCC 378
QY 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLeuIleLeuSerGly 120
DB 379 GGACAGAGCAGAAATTAACCTCCCAAGGAGGAGACTGATACATGCAAAATTTCTGTGACGG 438
QY 121 ThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
DB 439 ACACATGAAGGGGTGACTACTGGAAGCCCAATCTTGTATTTGTCCTCCAAACACAGATCAA 498
QY 141 IleGlySerAspHisArgGluIleAlaLeuValTyrArgProSerHisAlaAspAlaThr 160
DB 499 ATAGGCAGTGATCACCGTGAATAGCAATGTGTACCGACCTTCTCATGTCAGACGCAACT 558
QY 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyValArgSerSerClyArgLys 180
DB 559 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGAGGTGGAGGTCTCTCGGGCGAATA 618
QY 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysIleLeuLeuLysCysGly 200
DB 619 ACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAATTTCTTAAGCTCAATGTGA 678
QY 201 LeuGluLeuLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
DB 679 TTAGAGATCTTCTGCTTTTCCAAAGTGCATCAGGTGTGTCTCCCAAGACACGCGTT 738
QY 221 AspTyrGlySerValThrLeuGlnIleGluSerAsnIleValArgCysProAspPro 240
DB 739 GATTATGGGTCTGTAACTTGGAAACAGATAGAGCAACATCGTTAGATGTCCTGATCCA 798
QY 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260

```

```

Db 799 GAGTACGACAGAGAAGATGATAGACGCAATCGACAGAGTACGAGTTCGAGGGATTCGGTC 858
QY 261 GlyGlyValIleThrCysValAlaAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
DB 859 GGTGGAGTGTATCATCGTGTAGAAACGTTCTCTCGCGGGTTCGGTTCCTCTGTGTTTC 918
QY 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
DB 919 GACAGCTCGAATCCGAACTGGCAAAAGCTATGCTTTCTATTCTCTGCGACCAACGGGTTTC 978
QY 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
DB 979 GAGATTGCGACGCGATTTCGCGCGGACCGACTTCACAGCAAGTGCAGCATAATGATGACTTT 1038
QY 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
DB 1039 TATATGGATAAGGCTGGAAGTGTGAGGACACGCACTAATCGCTCGGGTGTGTGACAGGA 1098
QY 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
DB 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAGAGTTGCTTTTAAGCCGACACCATCTATC 1158
QY 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
DB 1159 GGGGTGAACAGAAACACCGTGTCAAGGCGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGG 1218
QY 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
DB 1219 CGCATGACCCATCGCTCGCCCTCGAGCTGTCTGTGTGGTAATCCATGGCGCGCTTG 1278
QY 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
DB 1279 GTCTCTATGACCACTGATGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338
QY 421 AlaLeuGlnGluProValGlySerPhe 429
DB 1339 GCACTTCAAGAACAGTTGGCTCTTTC 1365

RESULT 2
CNSOACWT 1488 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTJS82G07 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX815746
VERSION BX815746.1 GI:42473095
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1488)
AUTHORS Castellani,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1488)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
COMMENT BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

```

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

#### FEATURES

Location/Qualifiers  
 1..1488  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTSL8ZG07"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1488)  
 /gene="At1g48850"

#### gene

#### ORIGIN

#### Alignment Scores:

Pred. No.: 1.27e-143 Length: 1488  
 Score: 1441.00 Matches: 268  
 Percent Similarity: 83.8% Conservative: 65  
 Best Local Similarity: 67.51% Mismatches: 63  
 Query Match: 65.50% Indels: 1  
 DB: 3 Gaps: 0

US-10-677-179-8 (1-429) x CNS0ACWT (1-1488)

QY	31	AlaSerSerLeuAArgPheAlaValHisArgCysAArgThrAlaAArgLeuGluValLys	50
DB	149	TCCTCTCCGCGTTCAGATCTCTCTCGTACCCAAACAGGAACAATTCAGATACAA	208
QY	51	AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyGlySerHisGly	70
DB	209	GCTACTGGAAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTTGGAGATCATCGA	268
QY	71	GlyGlyValGlyCysValIleSerGlyCysProAArgIleProLeuThrGluAlaAsp	90
DB	269	GGAGGAGTGGTGTATCATGATGTTGTCTCTCTCTGATTCACATTCATGATCTGAT	328
QY	91	LeuGlnValGluLeuAspAArgArgProGlyGlnSerAArgIleThrSerThrArgLys	110
DB	329	TTGCAATTCGATCTCGATAGAAGGAGGCTGTGTACAGACGATCACAACTCTAGAAA	388
QY	111	GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro	130
DB	389	GAGACTGATATCTGCGGATATCGTCTGGAGTCTCTGAAGGAATGACACAGGAACCT	448
QY	131	IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150
DB	449	ATCCATGTTGTTGATACCAACACAGATCAGAGAGACTGTATGATGAAATGTCGTT	508
QY	151	ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValAlaVal	170
DB	509	GCCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAGTATGGTGTGATCAGT	568
QY	171	GlnGlyGlyCysArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu	190
DB	569	CAGGTTGAGGAAGATCTTCAGTACAGAGACCAATTTGGAAGATGTTCTCTGAGCTTG	628
QY	191	ProLysIleLeuLysCysGlyLeuLysGluIleLeuSerPheValSerLysVal	210
DB	629	GCDAAGAAATTTGAAGCAATTTGCAGAACTGATGATCTGTATGTTCCAGATT	688
QY	211	HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle	230
DB	689	CACCATGTTGTTCTCCAGAAAGATTTGTTAGACACGAGAAATTTAACTACGACAGATA	748
QY	231	GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle	250
DB	749	GAATAAATCATTTGTGATGTCCTTAATCCCGAGTATGCGGAAGATGATGATGTCGATT	808

QY	251	AspArgValArgVal-ArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAs	270
DB	809	GATGCTGTACAGACAAAGGAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	868
QY	270	nValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAl	290
DB	869	TGCTCCAGCTGGCTTGGTACACCGTTTTCGATAAATCTGAAGCAGAACTGGCAAAAGC	928
QY	290	aMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAs	310
DB	929	TTGATGTGCTGCTACCTGCAACTTAAGGGAATTTGAGTTTGGAAAGCGGCTTTGCGAGTACCT	988
QY	310	pLeuThrGlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgTh	330
DB	989	TTTGACTGCTTGTGACACCAATGATGATTTCTATCCGATGAAATGGAAGATACGATAC	1048
QY	330	rArgThrAsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPhe	350
DB	1049	CAGAACCAACCGATCTGGTGAATTCAGGGAGGATCTCAATGCTGAATAATAACAT	1108
QY	350	elysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGl	370
DB	1109	GAGAGTAGCTTCAAGCCACATCAACAAATTTGGAAGGAAGCAAAATACGTAACACAGAA	1168
QY	370	uArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAl	390
DB	1169	CAAGTAGAACCAACGAAATGATTCGCGTGTGCTGATGATCTGTGTGTGTGTGTGTGT	1228
QY	390	aValProValValGluSerMetAlaLeuValLeuMetAspGlnLeuMetAlaHisVa	410
DB	1229	TGTGCCAATGTTGGAAGCAATGTTGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1288
QY	410	lAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal	426
DB	1289	CGCAATGCTCAATTTGTTTCCAAATAATCCAGAGTTTCAGGAGAACCTCTC	1337

#### RESULT 3

#### CNS0ACWM

#### LOCUS

#### DEFINITION

CNS0ACWM 1494 bp mRNA linear HTC 06-FEB-2004  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSLSL8ZB05 of Silique of strain col-0 of Arabidopsis thaliana  
 (thale cress).

#### ACCESSION

VERSION BX818532.1 GI:42473287

#### KEYWORDS

HTC; GSLT cDNA.

#### SOURCE

Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

#### REFERENCE

#### AUTHORS

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

Unpublished  
 2 (bases 1 to 1494)  
 Genoscope.  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

#### COMMENT

The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein



length  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.  
 Location/Qualifiers  
 1. .1538  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTSL102H06"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCWSPORT 6"  
 complement(1. .1538)  
 /gene="Atlg48850"

gene

ORIGIN

Alignment Scores:  
 Pred. No.: 5.59e-139 Length: 1538  
 Score: 1398.00 Matches: 267  
 Percent Similarity: 83.17% Conservative: 64  
 Best Local Similarity: 67.09% Mismatches: 65  
 Query Match: 63.55% Indels: 3  
 DB: 3 Gaps: 0

US-10-677-179-8 (1-429) x CNSOACYT (1-1538)

QY	31	AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaAArgLeuGluValLys	50
Db	222	TCCTTCCTCCGCCGCTCAGATCTCTCTCGTACCCAAACAGGAAGAACTTCCAGATACAA	281
QY	51	AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyrglyGluSerHisGly	70
Db	282	GCTACTGGAAGTTCATATGGGACTCATTTTCGAGTTTCACCTTTTGGAGAAATCAGTGA	341
QY	71	GlyglyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp	90
Db	342	CGAGGAGTTGGTTGTATCATTCATTCATTCCTCTCTCGTATTCACCTTACTGAATCTGAT	401
QY	91	LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys	110
Db	402	TTTCAATTCGATCTCGATGAAGGAGGCTCGTCAGAGCAGGATCACAACTCTCTAGAAA	461
QY	111	GluThrAspThrCysLeuIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro	130
Db	462	GADACTGATACTTCCGGATATCGTCTCGAGTCACTGAAGGAATGACGACAGGAACACCT	521
QY	131	IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150
Db	522	ATCCATGTGTTGTATACCAACACAGATCAGAGAGGACTTGATTACAGTGAATATGTCGGT	581
QY	151	ValTyArgProSerHisAlaAspAlaThrTyArgPheLysTyrglyValArgAlaVal	170
Db	582	GCTATAGCCATCCATGCTGTATGCACTTATGACATGAAGTATGTGTCAATCAGT	641
QY	171	GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu	190
Db	642	CADGCTGAGGAAGATCTTCAGTAGAGACCACTTGGAGAGTGTCTCTCGAGCTTG	701
QY	191	ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal	210
Db	702	GCCAAGAAATT-TTGAAGCAATTTGCGAGCACTCAGATTTCTCGCTATGTCTCGCAATG	760
QY	211	HisGln-ValValLeuProGluAspAlaValAspTyrglySerValThrLeuGluGlnI	230
Db	761	CACCTTGTGTATCTCCAGAGAAGATTGGTAGACCACGAGAAATTTTACTTCGACGAT	820
QY	230	eGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaI	250
Db	821	AGAAATAACATTTGTAGATGCCCTTAATCCGAGTATGCCGAAGAAATGATAGTCCGAT	880
QY	250	eAspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAs	270
Db	881	TGATCTCTCAGGACAAAGGAACCTCTGTGGTGGTGTGTGACCTGCATTTTCGGA	940

of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from 1cm crown sections after 30 days of cold acclimation. The second is from 1cm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

## ORIGIN

## Alignment Scores:

Pred. No.: 5,41e-116 Length: 1160  
Score: 1182.50 Matches: 245  
Percent Similarity: 76.92% Conservative: 35  
Best Local Similarity: 67.31% Mismatches: 53  
Query Match: 53.75% Indels: 32  
DB: 7 Gaps: 5

US-10-677-179-8 (1-429) x CK210372 (1-1160)

```

Qy 2 ThrThrValProLysProGlnGln----- 9
Db 23 ACGCGTCGCGGAAACCAACCAATCTACCACCCCTCTCCAAATGGCAACGCGCCC 82
Qy 10 ValAlaHisSerArgAlaArgLeuAlaPro-----ArgAlaIleGly 23
Db 83 ACGTCGACACAGCTCGCGCGCGCGCCCTGGAGCTCCCTCCCGCGGGGTTCCGG 142
Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHisArgCysArg 43
Db 143 GCGTCTACGGACTCGCGCGCGCGC-----TCCGTCGCTTCTCCGTCGGTCGCGCGC 196
Qy 44 ThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAla 63
Db 197 GCCCGCGCGCTAGAGGTGAAGACGCTGGAACATCTTCGGGGACTACTTCCAGGTGCA 256
Qy 64 ThrTyrGlyGluSerHisGlyGlyValGlyCysValLysSerGlyCysProProArg 83
Db 257 ACTTATGAGAGTCTCATGGAGCGGTGTGGCTGCGTTATCATGTGTGTCCACCCAGA 316
Qy 84 IleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSer 103
Db 317 ATCCCACTCAGGAGGAAGACATG-----CAGAGC 346
Qy 104 ArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGlu 123
Db 347 AGAATAACACCCCAAGGAGGAGACCATCTTGTAAATCCCTTTCAGGACATACGAA 406
Qy 124 GlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySer 143
Db 407 GGAATGACCATCGGACACCAATTCATGTTCTTGTCCCAACACGATCAAGAGGGGT 466
Qy 144 AspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPhe 163
Db 467 GATTACACTGAATGGCTAAGCGGTACAGACCTTCCCATCGGATTTAACTTATGACCTC 526
Qy 164 LysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGly 183
Db 527 AAATACGGGTGTAGATCTGTGACGGAGGTGGAAGGTCATCGGAAGAGAAACCATTTGA 586
Qy 184 ArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysGlyLeuGluIle 203
Db 587 AGGCTAGCTGACGAGCTGTGGCAAGAAATTTTAAAGCTCAATGTGGGGTAGAATT 646
Qy 204 LeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGly 223
Db 647 CTAGCAATTTGTTTCCAAAGTGCATCAAGTGTGTTCTCTGAAGACGAGTTGATTACGAA 706

```

```

Qy 224 SerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAla 243
Db 707 GCTCTTACCCTGGATCAGATAGAGAACATCTGTAGATGTCTGTATCCAGAAATATGCA 766
Qy 244 GluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyVal 263
Db 767 CAGAGATGATTCATGCAATTCATTAAGTACGAATTAATGGGAATTCGATTTGGGGTG 826
Qy 264 IleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeu 283
Db 827 GTCACATGCATTCGCCAAATGCTCCTCGTGGGCTTCTCTCTGTATTGCAAACTG 886
Qy 284 GluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGly 303
Db 887 TGAGCTCTACTGGCAAGGCTATGCTTCTCTCTGCAAGAGGGGTTGAGATCCGT 946
Qy 304 SerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAsp 323
Db 947 CATGGAAATTCAGGTACTGACCTCTGAGTTGAGAT-AACGATAAGTTCTATATGGAT 1005
Qy 324 LysAlaGlySerValArgThrArgThrAsnArgSerGlyValGlnGlyIleSer 343
Db 1006 TAGGTTGAATGCTGAACACCAACCAACCATCG---CTCGCGGGGTTTCAGGAGGATATCA 1062
Qy 344 AsnValGluIle 347
Db 1063 AATGGGGAACATA 1074

RESULT 6
LOCUS CO082566
DEFINITION GR_Ea47E01.r GR_Ea Gossypium raimondii cDNA clone GR_Ea47E01 3', mRNA sequence.
ACCESSION CO082566
VERSION CO082566.1 GI:48752047
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
REFERENCE 1 (bases 1 to 923)
AUTHORS Kim H. Yu, Y. Kudrna, D. Hatfield, J., Stum, D., Mueller, C., Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and Wing, R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 47 row: E column: 01.
FEATURES
source
1..923
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea47E01"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/notes="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
Alignment Scores:

```



```
Pred. No.: 8.6e-115 Length: 923
Score: 1170.00 Matches: 221
Percent Similarity: 87.30% Conservative: 47
Best Local Similarity: 71.9% Mismatches: 39
Query Match: 53.18% Indels: 1
DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x C0082566 (1-923)

QY 112 ThrAspThrCysLeuValLeuSerGlyThrHisGluGlyValThrGlyThrProIleu 131
Db 2 ACAGACACATGCCGAATATATCTGACCTTCTGAAGGAGTGACTACTGGACACCAATC 61
QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
Db 62 CATGTACTGTGACCAAACTACTGATCAGAGAGACATGATTACAGGAATGTCAATAGCT 121
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
Db 122 TATAGGCCATCTCATGCTGATGCCACTTATGACATGAAATATGGTGTGAGGCGAGTGCAG 181
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
Db 182 GGTGTGTGTAGATCATCGGCCAGAGAAACCATTTGGAGAGATTGCTCTGGAGCTATTGCT 241
QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db 242 AAGAAATCTCAAGCTGTTTCAGAACTGAGGTCTTGCTATGTCTCTCAAGTTTCAAC 301
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
Db 302 CAGGTGTGTTCTCCAGATGGCTCAGTGCACATGACACTGTAACTCTTGTATCAGATAGAG 361
QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
Db 362 AGCAATATGTGAGGTGCGCAAACTCTGATACGCTGAGAAATGATTGCTGCTATTGAT 421
QY 252 ArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnVal 271
Db 422 GCTGTGCGGACNAGAGGAATCTATTGCGGAGTTGTGCATGATAGTGGAGGATGCT 481
QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
Db 482 CCAGTGGCTTGGTTTACCAGTTTGTATGATGCTTTGAAGCGGAGCTTGTCTAAGGCTGTA 541
QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
Db 542 ATGTCACCTACCTGCAACCAAGGCTTGAATTTGGGAGTGGATTGTCAGGCTACTTTGTTG 601
QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
Db 602 ACCGGTAGTGAACATATATGATGAGTTCTATCTGATGAACATGGAAGAAATAGGACAGA 661
QY 332 ThrAsnArgSerGlyGlyValGlnGlyLysSerAsnValGluIleValHisPheLys 351
Db 662 ACNAATCGCTCCGGTGGGATACAGGTTGGAATATCCATGCGGAAATATATAATATGAGA 721
QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
Db 722 GTAGCTTTCAAGGCCAACAGCTACAAATTTGGTGAAGAAACAGCACACAGTGAAGTGAAGA 781
QY 372 GlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaVal 391
Db 782 GAAGAGATAGAACTACTAGCCCGCGGTGCTCATGATCTCTTGTTGTCTCCCTCGAGCGGTG 841
QY 392 ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
Db 842 CCGATGGTTGAAGCGATGTAGCGTTGGTGTGCTTGTGTTGGACCAATA-ATGGCAATATGCA 900
QY 412 GlnCysGluMetPheAlaLeu 418
Db 901 CAATGTAACTCTTTCCTCTATC 921
```

RESULT 7  
CB654761  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CB654761 763 bp mRNA linear EST 09-APR-2003  
OSJNEC07113.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEC07113 5', mRNA sequence.  
CB654761  
EST.  
CB654761.1 GI:29658486  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 763)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aaa cga cgg cca gtg  
BACKWARD: gga aac agc tat gac cat g  
Plate: 07 row: 1 column: 13  
Seq primer: gta aaa cga cgg cca gtg.

FEATURES  
source  
1..763  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEC07113"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNEC"  
/note="vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ALIGNMENT SCORES:  
Pred. No.: 3.78e-111 Length: 763  
Score: 1135.00 Matches: 214  
Percent Similarity: 93.70% Conservative: 24  
Best Local Similarity: 84.25% Mismatches: 16  
Query Match: 51.59% Indels: 0  
DB: 6 Gaps: 0  
US-10-677-179-8 (1-429) x CB654761 (1-763)

QY 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeu 132  
Db 1 GACACTTGCAAAATTTCTTTCAGGACACATGAGGAATGACCACTGGGACACCAATTCAT 60  
QY 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152  
Db 61 GTTTTGTCTCCGAACACAGATCAGAGAGGGGTGATTACAGTGAATGGCTTAAGGCTTAC 120  
QY 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172  
Db 121 AGACCTTCACATGCAGATGCAACTTATGACTTCAATACGGTGTGTAGAGCAGTGCAGGA 180  
QY 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192  
Db 181 GGTGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGAGGCTCTTTCGCAAG 240



```

Qy 193 LysileLeuLysLeuLysCysGlyLeuGluLeuSerPheValSerLysValHisGln 212
Db 241 AAAATTCCTTAAGCTCAAACTGGAGTAGAGATCTGGCGTTTGTCGAAGTGCATCAA 300
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db 301 GTTGACTACAGAGATGCCGTTGATTCACACTGTAAACATGGAACAGATAGAAAGC 360
Qy 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db 361 AACATTGTTAGATGCTCTGATCCAGATATGCACAGAGATGATTCGCAATCGATAAA 420
Qy 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnValPro 272
Db 421 GTACGAGTTAGAGGTGATTCGATTGGTGTGTGCATCATGTCACGAAGAAATGTTCT 480
Qy 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetIleu 292
Db 481 CGTGGGATTGCTCTCTGTTATTCACAACTTGAGGCTGAATGGCGAAGCTATGCTT 540
Qy 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 541 TCTCTTCCTGCAAGCAAGGGTTTGAGATCGGCAGTGGATTTCGAGGTACTGACTACACT 600
Qy 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 601 GGAAGTGACATATGATGAGTTCATATGATGAGCTGGAATGTGAACCAAGAACT 660
Qy 333 AsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
Db 661 AATGTTTCAGCGGTGTGCAGGAGGATATCAAAATGTTGAAATATATATCTTCAGATGA 720
Qy 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThr 366
Db 721 GCTTTCAGGCCAACCGCACTATTTCGGAAGAAACCAACATACT 762

RESULT 8
CV256756/c
LOCUS CV256756 911 bp mRNA linear EST 22-SEP-2004
DEFINITION WS0244.B21_L13 PTX-D-ICC-N-A-14 Populus balsamifera subsp. trichocarpa x Populus deltoides cDNA clone WS0244_L13 3', mRNA sequence.
ACCESSION CV256756
VERSION CV256756.1 GI:52509731
KEYWORDS EST.
SOURCE Populus balsamifera subsp. trichocarpa x Populus deltoides
ORGANISM Populus balsamifera subsp. trichocarpa x Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 911)
AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakoff,R., Brown-John,M., Chand,S., Featherstone,R., Maeson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-6097
Email: bohlmann@interchange.ubc.ca
Plate: WS0244 row: L column: 13
High quality sequence stop: 911
POLYA=Yes.
FEATURES
Location/Qualifiers
source 1..911

```

```

/organism="Populus balsamifera subsp. trichocarpa x
Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"
/clone="WS0244_L13"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-ICC-N-A-14"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Cultured cells [de Sa MM et al. (1992) Plant Physiology 98:728-737] were grown in media (45mM) supplemented with either 50uM salicylic acid, 50uM benzothiadiazole, 50uM methyl jasmonate, 20ug chitosan or 200ul of Pollacia radiosa extract. Cells were harvested after a 3 hour treatment, along with untreated control cells. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.P. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."
Alignment Scores:
Pred. No.: 1,76e-109 Length: 911
Score: 1120.50 Matches: 215
Percent Similarity: 78.83% Conservative: 42
Best Local Similarity: 65.95% Mismatches: 39
Query Match: 50.93% Indels: 30
DB: 7 Gaps: 1

US-10-677-179-8 (1-429) x CV256756 (1-911)
Qy 48 GluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGlu 67
Db 899 AAAATAAATCGCGGGAAGCACATTTGGGACTAAATTTAGAGTAACATACATTTGGTGAG 840
Qy 68 SerHisGlyGlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThr 87
Db 839 TCACACGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
Qy 88 GluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSer 107
Db 779 GAAGCTGATATGCAATTTGATAGTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Qy 108 ThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThr 127
Db 719 CCGAGAAAGAGAGCGGATACCTGCATAATATCTTCGGTGTGTTCTGAAGGAGCTGACTACT 660
Qy 128 GlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGlu 147
Db 659 GGGACGCCGATTCATGTATTGTTGATCAAAATACCTGATCAGAGAGGAGGATTAGATTACAGTGAA 600
Qy 148 IleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyVal 167
Db 599 ATGTCAGTTGCTTACAGCCCTTTCATGTCAGACGCAACTTATGACATGAGTATGGTGT 540
Qy 168 ArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAla 187
Db 539 AGATCAGTTTCAGGTGGGGGTAGATCTTCAGCAAGAGAAACAATTCGAAGAGTGTGCT 480
Qy 188 GlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluLeuSerPheVal 207
Db 479 GGAGGTGTTGCTAAGAAAAATTCACAACTATATGACAGGAACCTGAGATTCTTGCTTATGTC 420

```

QY 208 SerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeu 227  
 Db 419 TCTCAAGTCACCAAGGTGTGTTCTCCAGAGGTGTGTTGATCAGCACTCTCTACACTT 360  
 QY 228 GluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIle 247  
 Db 359 GATCAGATGGAGCAATATTGTTCAGATGTCAGATCCCTGAATATGTCAGAGAAGATGATA 300  
 QY 248 AspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThrCysVal 267  
 Db 299 GCTGCCATTGATGCTGTCGAGTGAAGAGGATCTCTGTGGTGT 255  
 QY 268 AlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeu 287  
 Db 255 ----- 255  
 QY 288 AlaLysAlaMetLeu-SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAl 307  
 Db 254 -----GTTGTCACTACTCTGCAACAAGAGGCTTTGAATTTGGGAGTGGATTGCT 207  
 QY 307 aGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySe 327  
 Db 206 AGCACTCTCTGACTGGGAGTGAACATAATGACGAGTCTTACACAGCAAAATCATGGAAG 147  
 QY 327 rValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluI 347  
 Db 146 AATCCGAACAAGACAAATCGTTCTGTGGAATACAGGAGGAGGATATCAATGGTGAAT 87  
 QY 347 eValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrVa 367  
 Db 86 TATAAATCAGATGAGTATGCTTCAAGCAACATCTACATAGGAAGAAAGCAACATACACT 27  
 QY 367 lSerArgGluArgGln 372  
 Db 26 TACTAGAGAAAAA 11

RESULT 9  
 LOCUS Col14396  
 DEFINITION GR\_EB015107.r GR\_Eb Gossypium raimondii cDNA clone GR\_Eb015107  
 3', mRNA sequence.  
 ACCESSION Col14396  
 VERSION Col14396.1 GI:48813083  
 KEYWORDS EST.  
 SOURCE Gossypium raimondii  
 ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 1. (bases 1 to 865)  
 Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,  
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
 Wing,R.A.  
 Global assembly of Cotton ESTs  
 Unpublished (2004)  
 Contact: Rod A. Wing  
 Arizona Genomics Institute  
 The University of Arizona  
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: http://genome.arizona.edu  
 Plated: 015 row: 1 column: 07.  
 Location/Qualifiers

FEATURES  
 source  
 1. .865  
 /organism="Gossypium raimondii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:29730"  
 /clone="GR\_EB015107"  
 /tissue\_type="floral"  
 /dev\_stage=" 3 to +3 DPA"  
 /lab\_host="DH10B"  
 /clone\_lib="GR\_Eb"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 6,53e-106 Length: 865  
 Score: 1087.00 Matches: 200  
 Percent Similarity: 88.0% Conservatives: 44  
 Best Local Similarity: 72.20% Mismatches: 32  
 Query Match: 49.41% Indels: 1  
 DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x Col14396 (1-865)

QY 48 GluValLysAlaSerGlyAsnThrPheGlyAsnTyr-PheGlnValAlaThrTyrGlyGl 67  
 Db 37 CAGATACAAAGCAGCTGGGAGCACATTTGGAAATATTCTTCGGTGTCAACAATTTGGAGA 96  
 QY 67 uSerHisGlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuTh 87  
 Db 97 ATCTAATGGAGTGTGTGTGTGTATAGTGTGATGATGCTCTCTAGGATTCCTCTCTC 156  
 QY 87 rGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSe 107  
 Db 157 AGAGCTGATTTGCAAGGTGATCTTGACCGAAGGAGCGGGTCAAGCCGTTACTACTAC 216  
 QY 107 rThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrTh 127  
 Db 217 CCCAAGGAAGAGACAGACATGCGGAATATATTTCTGGAGTTTCTGAAGGAGTACTAC 276  
 QY 127 rGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGl 147  
 Db 277 TGGAAACCAATACATGATCTGTACCAAAATCTGATCAGAGAACCAAGATTACAAGGA 336  
 QY 147 uIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyVa 167  
 Db 337 AATGTCATAGTATTAGGCCATCTCATGCTGTATGTCACCACTATGACATGAATATGGTGT 396  
 QY 167 lArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAla 187  
 Db 397 CAGGCGAGTGCAGGGTGTGGCAGATCATCAGCCAGAGAACCAATTTGAAGAGTTGCCTC 456  
 QY 187 aGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheVa 207  
 Db 457 TGGAGCTATTGCTAAGAAAAATTTCTCAAGCAATTTTTCAGAACTGAGGTTCTTGCATATGT 516  
 QY 207 lSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLe 227  
 Db 517 CTCTCAAGTTCAACAGGTTCTTCTTCCAGATGGGTGTCAGTACCATGACATGTAACCTCT 576  
 QY 227 uGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetI 247  
 Db 577 TGACCAGATAGAGCAATATTGTAAGGTGCCAAATCTCTGATACCTGAGAAATGAT 636  
 QY 247 eAspAlaIleAspArgValArgValArgLysSerValGlyValIleThrCysVa 267  
 Db 637 CGCTCTATCGATGCTGTGCGGATAAGAGGAGATTTCCGTTCGCGGTGTGTGCACGTGCA 696  
 QY 267 lAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLe 287  
 Db 697 AGTGAGGAATGCTCCGCGTGGGCTTGGTTTCCAGTATTTTGAATAGCTTTGAAGCAGAGCT 756  
 QY 287 uAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAl 307  
 Db 757 TGCTAAGGCTGTAATGTCAGTACCTGTCATCCAAAGGCTTTGAAGTGGAGTGGATTTTC 816  
 QY 307 aGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAsp 323  
 Db 817 AGGTGCTTTTCTAACCGGTAGTGAACATAATGACGAGTCTCTATACTAGT 865

RESULT 10

```

BG600177      BG600177      786 bp      mRNA      linear      EST 07-MAR-2003
LOCUS          EST505072 csts Solanum tuberosum cDNA clone cSTS28C6 5' sequence,
DEFINITION     mRNA sequence.
ACCESSION      BG600177
VERSION        BG600177.1  GI:13617313
KEYWORDS       EST.
SOURCE         Solanum tuberosum (potato)
ORGANISM       Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE      1 (bases 1 to 786)
AUTHORS        van der Hoeven R., Bezzerides J., Sun, H., Cho, J., Chiemiango, A.,
                Bougri, O., Buell, C.R., Roming, C., Tanksley, S. and Baker, B.
TITLE          Generations of ESTs from sprouting potato eyes
JOURNAL        Unpublished (2000)
COMMENT        Contact: Robin Buell
                The Institute for Genomic Research
                9712 Medical Center Dr, Rockville, MD 20850, USA
                Email: potato-array@igr.org
                This clone can be obtained from the University of Arizona Genomics
                Institute. Orders can be made through URL:
                http://genome.arizona.edu/orders/
                Seq primer: M13P-R.

FEATURES       Location/Qualifiers
source         1..786
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Kennebec"
                /db_xref="taxon:4113"
                /clone="cSTS28C6"
                /tissue_type="sprouting eyes from tubers"
                /dev_stage="12-14 weeks post harvest"
                /lab_host="SOLR"
                /clone_lib="cSTS"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                taken from tubers. The tubers were incubated at 26C in the
                dark for 2-3 weeks prior to sprouting. The eyes were
                frozen in liquid nitrogen immediately upon removal from
                tubers."

ORIGIN
Alignment Scores:
Pred. No.:      3.9e-101      Length:      786
Score:          1042.00      Matches:    197
Percent Similarity: 89.4%      Conservative: 33
Best Local Similarity: 76.6%      Mismatches: 26
Query Match:    47.36%      Indels:     1
DB:              4              Gaps:       0

US-10-677-179-8 (1-429) x BG600177 (1-786)

Qy      46 ArgLeuGluValLysAlaSerGlyValThrPheGlyAsnTyrPheGlnValAlaThrTyr 65
Db      13 CGACTAGAGATCAGACGGCTGCTGGTAATACATTTGGAAATATTTCCTGGGTACACTTTT 72

Qy      66 GlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProProArgIlePro 85
Db      73 GGAGAAATCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 132

Qy      86 LeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIle 105
Db      133 CTCTCAGAGTCGGATATCGAAGTGAACCTTGACAGGAGGAGCCAGGTCAAAGCCGAATT 192

Qy      106 ThrSerThrArgLeuGluThrAspThrCysValIleLeuSerGlyThrHisGluGlyVal 125
Db      193 ACCACCAAGAAAAGAGACTGATCTGCAAAATATCATCAGGACATGCCGATGGTTG 252

Qy      126 ThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHis 145
Db      253 ACTACGGGATCTCCCAATCAAGTTGAAGTACCACACTGACCAGAGAGGAATGACTAT 312

```

```

Qy      146 ArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyr 165
Db      313 AGTGAATAGTCCCTTGTCTTATAGCCACTCTCATGCTGATCAACTTATGATCTTCAAGTAT 372

Qy      166 GlyValArgAlaValGlnGlyGlyValArgSerSerGlyArgLysThrValGlyArgVal 185
Db      373 GGAGTTAGATCAGTACAGGGGGGGGGGTAGATCTTCAGCAAGAGAAACCATCGGAGAGTC 432

Qy      186 AlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSer 205
Db      433 GCTGCTGGAGCTGTTCAGCAAGAAAATCTCAAACTTTATTTCAGGAACCTGAGATCTTGTCT 492

Qy      206 PheValSerLysValHisGlnValValLeuProGluAspAlaValAlaSerGlySerVal 225
Db      493 TATGTTTTCAAGTTCACATGTTGTACTTCCAGAGGATTTGGTTGATAACCGAGCTGTG 552

Qy      226 ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys 245
Db      553 ACCTAGAGCAGATTGAAGCAATATTGTTGTCGCCGACCTGATATGCGAAGAAAG 612

Qy      246 MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThr 265
Db      613 ATGATTGCTGCCATTGATTATGTACGAGTACGAGTACGAGGGGATTCCTGTTGGTGTCTGTA 672

Qy      266 CysValAlaArgAsnValProArg-GlyLeuGlySerProValPheAspLysLeuGluSe 285
Db      673 TGCATTGTTAGAAATGTTCCACGAAGGTCTTGGTACACCATCTTTCGATAAATTTGAAGC 732

Qy      285 xGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGlu 301
Db      733 TGAGCTGGCTAAGCTTGCATGTCATTCACGAGCACCACCAAGGGTTTTCAG 781

RESULT 11
BM411032      835 bp      mRNA      linear      EST 22-JAN-2002
LOCUS          EST585359 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION     CLEGS5M18 5' end, mRNA sequence.
ACCESSION      BM411032
VERSION        BM411032.1  GI:18262662
KEYWORDS       EST.
SOURCE         Lycopersicon esculentum (tomato)
ORGANISM       Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 (bases 1 to 835)
AUTHORS        Alcalá, J., Vrebalov, J., White, R., Vision, T., Karameycheva, S.A.,
                Tsai, J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S.,
                Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
                Giovannoni, J.
TITLE          Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL        Unpublished (2002)
COMMENT        Contact: CUGI
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson, SC 29634, USA
                Email: http://www.genome.clemson.edu/orders/index.html
                This clone is available through the Clemson University Genomics
                Institute
                Seq primer: T3.
                Location/Qualifiers
                1..835
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultivar="TA496"
                /db_xref="taxon:4081"
                /clone="cLEGS5M18"
                /tissue_type="pericarp"
                /dev_stage="breaker"
                /lab_host="SOLR"
                /note="Vector: pBluescriptSKmCvadapt; Site_1: EcoRI;

```



Db 363 TTGAGATCGGTAGTGGATTTTCAGGTACTGACCTAACTGCAAGTGCATACAGATGAG 422  
Qy 320 PheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGln 339  
Db 423 TTTATATGGAGCGGTGGAAATGTAAGAACACAGCAACATCGCTCGGGCGGTACAG 482  
Qy 340 GlyGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSer 359  
Db 483 GGAGGGATATCAATGGTGAACACTATATATCTCAAGTAGCTTTCAGCCACAGCAACT 542  
Qy 360 IleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArg 379  
Db 543 ATTTGGGAAGAGCAAAATACTGTAAACAGGATCATGAGGATATCGAACTTCTGACAAAG 602  
Qy 380 GlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAla 399  
Db 603 GGTGCCCATGACCCATGTGTTCCTCGGGCTGTTCCAATGGTGGAGCATGGCTGCA 662  
Qy 400 LeuValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMet 415  
Db 663 TTGTTCTCATGGACCACTGATGCCACATGTTGCTCAGTCGAGATG 710

RESULT 13  
CN144654  
LOCUS  
DEFINITION WOUND1\_23\_E02\_g1\_A002 Wounded leaves Sorghum bicolor CDNA clone  
ACCESSION WOUND1\_23\_E02\_A002 5', mRNA sequence.  
VERSION CN144654.1 GI:45983326  
KEYWORDS EST.  
SOURCE Sorghum bicolor (sorghum)  
ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCD  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 814)  
Cordonnier-Pratt M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,  
Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M.,  
Anfuso, C., Chhabra, D., Johnson, H., Kamran, D. and Pratt, L.H.  
A Sorghum EST database: mechanically damaged and methyl  
jasmonate-treated leaves  
Unpublished (2003)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

Librarty constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
the Human Genome Center, University of Tokyo Institute of Medical  
Science; plant material and RNA prepared at Texas A & M University;  
sequencing done in the Laboratory for Genomics and Bioinformatics,  
University of Georgia. Sequence ends have been trimmed to exclude  
vector and regions below Phred quality 16. Three-prime sequences  
are presented as their reverse complement and have been trimmed to  
exclude polyA.  
seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).  
Location/Qualifiers  
1..814  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="WOUND1\_23\_E02\_A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="wounded leaves"  
/note="Organ: Leaf; Vector: pME18S-FL3; Site 1: XhoI;  
Site 2: XhoI; The library was prepared from polyA+ RNA  
harvested from 8-day-old hydroponically grown, BTx623  
sorghum seedlings. For some plants, one-half of the second  
leaf was crushed without damaging the midvein. For others,  
methyl jasmonate was added to the growth medium to a final

FEATURES  
source

concentration of 100 uM. Leaves were harvested 3 and 27 hr  
after treatment and pooled. Double-stranded cDNA was  
cloned unidirectionally into different DraIII sites of the  
pME18S-FL3 vector (5-prime DraIII site is CACTGTGG,  
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA  
insert."

## ORIGIN

Alignment Scores: 1.49e-99 Length: 814  
Pred. No.: 1027.50 Matches: 204  
Score: 86.96% Conservative: 16  
Percent Similarity: 80.63% Mismatches: 30  
Best Local Similarity: 46.70% Indels: 3  
Query Match: 7 Gaps: 2  
DB:

US-10-677-179-8 (1-429) x CN144654 (1-814)

Qy 4 ValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArgAlaIleGly 23  
Db 62 GTGTGGCGCGCGGTGGCGCGCAGGGCGTCCACGGGTCTCTCCCGCGGGGATAGGC 121  
Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaVal---HisArgCys 42  
Db 122 GCGTCCCGGAGTCCGCGCGCGCGG-----TCGCTCGCTTCTCCGTCGCGCGCGCGC 175  
Qy 43 ArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnVal 62  
Db 176 CGCGCGCTCGCTAGAGGTGAAGGCATCGGGAATGTGTTCGGGAATCTACTTCCAGGTT 235  
Qy 63 AlaThrTyrGlyGluSerHisGlyGlyValGlyCysValLysSerGlyCysProPro 82  
Db 236 GCACCTATGGCAATCCCATGGAGGGGTGTGTGTGTATCATGGTGGTCCCAACC 295  
Qy 83 ArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGln 102  
Db 296 AGAATTCCTCTCTGAGGCAGACATCAAGTAGTAGAAGTTCGATAGAGACGTCGGGTCAA 355  
Qy 103 SerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHis 122  
Db 356 AGTAGAATAACAACCCCAAGAGGAGACTGATACATGCAAAATTTCTATCAGGACACAT 415  
Qy 123 GluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGly 142  
Db 416 GATGGGATGACTACTGCAACCAATTCAGCTCTTTGTCCGGAACACAGATCAAGAGGC 475  
Qy 143 SerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAsp 162  
Db 476 GGTGATTACAGTGAATGGCTAAGGCGTACAGACCATCCCATGCAGATGCAACCTATGAC 535  
Qy 163 PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal 182  
Db 536 TTCAAGTATGTTGTGACAGAGTGTGACGGAGGGGGAAGGTTCATGGCCAGAGAACCAT 595  
Qy 183 GlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGlu 202  
Db 596 GGCAGGGTGGCTGCGAGGAGCTCTTGCAAAAGAAATTTCTGAAGCTCAATCAGAGTGGAG 655  
Qy 203 IleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyr 222  
Db 656 ATCTTGGCAATTTGTTCTTAAAGTGCATCAAGTCGTACTTCCAGAGATGCAAGTTGATTAT 715  
Qy 223 GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr 242  
Db 716 GAGACTGTAGCTTGGAAACAGATAGAGCAACATTTGTAGATGCTCTCTGATCCAGAAATAT 775  
Qy 243 AlaGluLysMetIleAspAlaIleAspArgValArgVal 255  
Db 776 GCAGAGAGATGATTGCTGCCATTGATAAAGTACAGATT 814

RESULT 14  
CF303160  
LOCUS

CF303160 723 bp mRNA linear EST 15-AUG-2003

DEFINITION ABFI-01-104.g1 ABF3-overexpressing transgenic rice lambda phage CDNA library (ABFI) Oryza sativa (japonica cultivar-group) cDNA clone ABFI--01-104, mRNA sequence.

ACCESSION CF303160.1 GI:33674921

VERSION EST...

KEYWORDS Oryza sativa (japonica cultivar-group)

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1. (bases 1 to 723)

AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University  
Yongin, Gyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

FEATURES

Location/Qualifiers

1..723

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nackdong"

/db\_xref="taxon:39947"

/clone="ABFI--01-104"

/tissue\_type="leaf"

/dev\_stage="14 days after germination"

/lab\_host="E.coli SOLR"

/clone\_lib="ABF3-overexpressing transgenic rice lambda phage cDNA library (ABFI)"

/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; Leaf was dried for 2hrs. cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was prepared from ABA-responsive element binding transcription factor 3 overexpression line."

## ORIGIN

Alignment Scores:

Pred. No.: 4, 89e-99 Length: 723

Score: 1022.00 Matches: 199

Percent Similarity: 93.62% Conservative: 21

Best Local Similarity: 84.68% Mismatches: 14

Query Match: 46.45% Indels: 2

DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x CF303160 (1-723)

QY 89 AlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThr 108

DB 20 GCAGATATGCAAGTAGAAGTCTGACCGGAGCGCCAGCCAGCAGAGATTAACACCCCA 79

QY 109 ArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGly 128

DB 80 AGAAAGGAGAGCTGACACTTGCAAAATCTTTCAGGGACACATGAAGGAATGACCACTGGG 139

QY 129 ThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIle 148

DB 140 ACACCAATTCATGTTTGTCCGACACACAGATCAGAGGGGGGTGATTACAGTGAATG 199

QY 149 AlaAsnValThrArgProSerHisAlaAspAlaThrThrAspPheLysThrGlyValArg 168

DB 200 GCTAAGGCTACAGACCTTCATGCAGATGCACTTATGACTTCAAAATACGGTGTAGA 259

QY 169 AlaValGlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGly 188

DB 260 GCAGTGCAGGAGGTGGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGA 319

QY 189 AlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSer 208

DB 320 GCTCTTGCAAGAAAATTTCTTAAGCTCAATCTCGAGTAGAGATCTTGCGCTTGTGTGCC 379

QY 209 LysValHisGlnValValLeuProGluAspAlaValAspThrGlySerValThrLeuGlu 228

DB 380 AAGGTGCATCAAGTTGTACTACCAAGATCCCGTTGATTATGACACTGTAAACAATGGAA 439

QY 229 GlnIleGluSerAsnIleValArgCysProAspProGluThrAlaGluLysMetIleAsp 248

DB 440 CAGATAGAAACAACATTTAGATGCTCTGATCAGAAATATGCACAGATGATTGAT 499

QY 249 AlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAla 268

DB 500 GCAATCGATAAAGTACGAGTTAGAGTTCGATTGGTGGTGGTGCACATGCAATGTC 559

QY 268 aArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAl 288

DB 560 AAGAAATGTTCTCTCGTGGGATGGCGTCTCTGTTATTTGACAAACTTGAGGCTGAATTGCC 619

QY 288 aLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaG1 308

DB 620 GAAGACTATGCTTCTCTCTCGCAGCAAG-GGGTTTGATCGGACGATGATTTGCAAG 678

QY 308 YThrAspLeuThrGlySerGluHisAsnAspGluPheThrMet 322

DB 679 TACTGACTACACTGGAAGTGAGCATATGATGAGTTCATATG 721

RESULT 15

CK279831

LOCUS

DEFINITION EST725909 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POAEE54 5' end, mRNA sequence.

ACCESSION CK279831

VERSION CK279831.1 GI:39836809

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum

REFERENCE 1. (bases 1 to 755)

AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

TITLE Generation of ESTs from abiotic stressed potato tissue

JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST725910

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..755

FEATURES

source

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POAEE54"

/tissue\_type="abiotic stress treated leaf and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="potato abiotic stress cDNA library"

/note="vector: pCMVSPORT6.1; site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.:	5.2e-99	Length:	755
Score:	1022.00	Matches:	188
Percent Similarity:	91.39%	Conservative:	35
Best Local Similarity:	77.05%	Mismatches:	21
Query Match:	46.45%	Indels:	0
DB:	7	Gaps:	0

US-10-677-179-8 (1-429) x CK279831 (1-755)

Qy	46	ArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyr	65
Db	22	CCCTAGAGATACAGCGTCTGGTATGATATTTGGAACTACTCCCGGTAACTATTT	81
Qy	66	GlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProArgIlePro	85
Db	82	GGAGAAATCTCATGGTGGTGGAGTTGGTTGTTATTATTTGATGGATGTCCCGGCTCCCA	141
Qy	86	LeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIle	105
Db	142	CTTCTGAATCTGATATGATGAGTGAACTTGACAGAGAGCCAGGTCAAAGCGGATT	201
Qy	106	ThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyVal	125
Db	202	ACCACCTTAGGAAAGAGACTGACACTTGCAAAATTTTCATCAGGCACCTGCAGATGGGCTG	261
Qy	126	ThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHis	145
Db	262	ACTACTGGATCTCCAAATCAAGTTGCAAGTACCTACACTGATCAGAGAGAAATGACTAT	321
Qy	146	ArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyr	165
Db	322	AGTGAATGTGCGTTTACAGGCCATCTCATGCAGATGCCACTTATGACTTCAAGTAT	381
Qy	166	GlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyVal	185
Db	382	GGAGTGAGATCTGTACAGGGGGTGTAGATCATCAGCAGAGACCATTTGGGAGATT	441
Qy	186	AlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSer	205
Db	442	GCTGCTGGAGCAGTTGCTTAAGAAATTTCTCAAACCTTATCTGGAGCTGAGGTTCTTGCT	501
Qy	206	PheValSerLysValHisGlnValValLeuProGluAlaValAspTyrGlySerVal	225
Db	502	TATGTTTCTCAAGTTTCAACAAGTTTGTACTTCTCGAGGATTTGATTGATCATCAGAAATGTG	561
Qy	226	ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys	245
Db	562	ACTTAGAGCAGATAGAAAGCAATATTTGCCATGCCAGATCCGGAAATATGCAGAGANG	621
Qy	246	MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThr	265
Db	622	ATGATTGCTGCTATTGATGCTGTACGAGTGAGAGGGGATTCTGTTGGTGGTGTGTAAC	681
Qy	266	CysValAlaAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSer	285
Db	682	TGCATTGTTAGAAATCTCCACGGGGTCTTGGTACACCAGTCTTCGATAAACTTGAAGCT	741

Qy	286	GluLeuAlaLys	289
Db	742	GAGTAGCCCA	753

Search completed: August 27, 2005, 07:09:56  
Job time : 3372 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 27, 2005, 04:27:24 ; Search time 223 seconds  
(without alignments)

3147.815 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQVQVAHSRRARLAPR.....VAQCEMFALNTALQEPVGSF 429

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool\_p/US10677179/runat\_25082005\_150240\_24862/app\_query.fasta\_1.583  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNIT9=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10677179 @CNG\_1\_1\_105 @runat\_25082005\_150240\_24862 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	4	US-09-743-207-7
2	1797.5	81.7	1635	4	Sequence 7, Appli
3	1257	57.1	966	4	Sequence 11, Appli
4	1066	48.5	1015	4	US-09-743-207-5
5	859	39.0	1134	4	US-09-489-039A-5375
6	850	38.6	1089	4	Sequence 5375, Ap
7	847	38.5	1239	3	US-09-543-681A-120
8	847	38.5	4530	3	US-09-064-693A-20
9	845	38.4	640681	4	US-09-064-693A-26
10	840.5	38.2	1830121	4	US-09-790-988-1
11	840.5	38.2	1830121	4	US-09-557-884-1
12	826.5	37.6	1080	4	US-09-643-990A-1
					US-09-248-796A-4099

13	800.5	36.4	1664976	4	US-08-916-421B-1	Sequence 1, Appli
14	800.5	36.4	1664976	4	US-09-692-570-1	Sequence 1, Appli
15	786	35.7	1140	4	US-09-252-991A-13588	Sequence 13588, A
16	786	35.7	1233	4	US-09-252-991A-13477	Sequence 13477, A
17	786	35.7	1263	4	US-09-252-991A-13857	Sequence 13857, A
18	756.5	34.4	6464	3	US-09-221-017B-168	Sequence 168, App
19	741	33.7	1104	4	US-09-328-352-3536	Sequence 3536, Ap
20	731	33.2	2312	4	US-09-103-331-1	Sequence 1, Appli
21	729	33.1	1125	4	US-09-540-236-310	Sequence 310, App
22	729	33.1	21410	4	US-09-596-002-12	Sequence 12, Appli
23	727	33.0	2312	4	US-09-631-594-45	Sequence 45, Appli
24	722.5	32.8	1053	3	US-09-610-040-1	Sequence 1, Appli
25	722.5	32.8	1053	3	US-09-610-040-7	Sequence 7, Appli
26	722.5	32.8	1053	4	US-10-267-763-1	Sequence 1, Appli
27	722.5	32.8	1053	4	US-10-267-763-7	Sequence 7, Appli
28	644	25.3	1837	4	US-09-103-331-3	Sequence 3, Appli
29	644	25.3	1837	4	US-09-631-594-54	Sequence 54, Appli
30	642.5	29.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli
31	642.5	29.2	1230230	4	US-09-438-185A-1	Sequence 1, Appli
32	614.5	27.9	17245	4	US-09-902-540-1073	Sequence 1073, Ap
33	611	27.8	1020	4	US-09-902-540-9186	Sequence 9186, Ap
34	467	21.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
35	467	21.2	441529	3	US-09-103-840A-1	Sequence 1, Appli
36	459.5	20.9	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
37	437	19.9	1278	3	US-09-134-001C-306	Sequence 306, App
38	437	19.9	2168	4	US-09-710-279-3562	Sequence 3562, Ap
39	426	19.4	5532	4	US-08-956-171E-530	Sequence 530, App
40	426	19.4	5532	4	US-08-781-986A-530	Sequence 530, App
41	424	19.3	541	4	US-09-743-207-13	Sequence 13, Appli
42	419.5	19.1	1176	4	US-09-107-433-2027	Sequence 2027, Ap
43	419.5	19.1	19390	3	US-08-961-527-86	Sequence 86, Appli
44	416.5	18.9	1167	4	US-09-583-110-882	Sequence 882, App
45	413.5	18.8	1188	4	US-09-134-000C-1577	Sequence 1577, Ap

#### ALIGNMENTS

RESULT 1  
US-09-743-207-7  
; Sequence 7, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Calico, Saverio E.  
; APPLICANT: Falco, Saverio O.  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207  
; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 1626  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-743-207-7

Alignment Scores:  
Pred. No.: 1.72e-240 Length: 1626  
Score: 2200.00 Matches: 429  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-677-179-8 (1-429) x US-09-743-207-7 (1-1626)

Qy 1 MetThrValProLysProGlnValAlaHisSerArgAlaLeuAlaProArg 20



Db 418 GATGGGATGACTACTGGTACACCAATTCACGTCTTTGTCCCAACACAGATCAAGGGGT 477  
QY 143 SerAspHisArgGluLeuAlaAenValTyrArgProSerHisAlaAaspAlaThrTyrAsp 162  
Db 478 GGTGATTACGTGAAATGCTTAAGGGGTACAGACCATCCCATCGATGCAACCTATGAC 537  
QY 163 PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal 182  
Db 538 TTCAGTATGAGCTTAGAGCTGTGACAGGAGGTGGAAGGTTCATCAGCCAGAGAAACCAAT 597  
QY 183 GlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysLysLysLysLys 202  
Db 598 GGCAGGGTGGCTGCAGAGAGCTTTCGCAAGAAATTCGAAAGCTCAAAATCAGAGTGGAG 657  
QY 203 IleLeuSerPheValSerLysValHisGlnValValLeuProLysAlaValAspTyr 222  
Db 658 AYTCTGGCATTTGTTCTTAAGTGCACCAAGTCTGTTCTCCAGAGATGAGTTGATTAAT 717  
QY 223 GlySerValThrLeuGluGlnIleGlySerAsnIleValArgCysProAspProGluTyr 242  
Db 718 GAGACTGTAACTTGGAAACATATAGAGAGCAACATCGTTAGATGCTCTGATCCAGATAT 777  
QY 243 AlaGluLysMetIleAaspAlaIleAaspArgValArgValArgGlyAspSerValGlyGly 262  
Db 778 GCAGAGAGATGATTCTGCTCATTTGATACGTACGATGAGTGTAGAGAGATTCATTAAT 837  
QY 263 ValIleThrCysValAlaArgAenValProArgGlyLeuGlySerProValPheAspLys 282  
Db 838 GTGCTCATATGATTCGCAAGAAATGTTCTCGTGGTCTTGGCTCTCTGTTTTCAGAAA 897  
QY 283 LeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIle 302  
Db 898 CTTGAAGCTGAATGCGCAAGAGCAATGCTTCTCTCTGCTGCAAGCAAGGGTTGAGATT 957  
QY 303 GlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMet 322  
Db 958 GGCAGTGGGTTCGCTGCTGACGACTTTACTGGAAGTGAGCATAAATGATGATTTCTATATG 1017  
QY 323 AspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIle 342  
Db 1018 GATGAGCGTGGAAATGTGAGGACACCACTTAATCGCTCAGGCGGTGTTTTCAGGAGGATA 1077  
QY 343 SerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362  
Db 1078 TCAATGGTGAATATTATTACTTCAAGTGGCTTTTAAAGCAACAGCACTATCGGAAG 1137  
QY 363 LysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuAlaArgGlyArgHis 382  
Db 1138 AAGCAAAATACTGTGTCAAGGGAGCATGAGGATGTTGAATCTTTGGCAAGGGGGGCCAT 1197  
QY 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402  
Db 1198 GACCCCTGTGTGCTCCTCGAGCTGTTCTATGTTGGAATCATGCTGCTGCTGCTGCTG 1257  
QY 403 MetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeu 422  
Db 1258 ATGGACCAGCTCATGGCGATATGCCCAGTGTGAGATGTTTCCGCTGACCTTCCCTTA 1317  
QY 423 GlnGluProValGlySer 428  
Db 1318 CAAGAGCCCATTTGGCTCT 1335

## RESULT 3

US-09-743-207-11  
; Sequence 11, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Fember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207

; CURRENT FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 11  
; LENGTH: 966  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-743-207-11

Alignment Scores:  
Pred. No.: 2,316-133 Length: 966  
Score: 1257.00 Matches: 245  
Percent Similarity: 89.30% Conservative: 22  
Best Local Similarity: 81.94% Mismatches: 26  
Query Match: 57.14% Indels: 6  
DB: 4 Gaps: 2

US-10-677-179-8 (1-429) x US-09-743-207-11 (1-966)

QY 15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro 30  
Db 75 GCGCGCTCTCCCG 134  
QY 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValLys 50  
Db 135 GCT-----TCCTCGCTTCTCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 188  
QY 51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70  
Db 189 GCGTCTGCAATGATTTGGGAACCTACTTCCAGGTTCACACTTATGGAGAGTCTCATGA 248  
QY 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90  
Db 249 GCGGTGTTGGTGGTAATCAGTGGATGCCCGCCAGCAATCCCACTTACTGAAGCAGAT 308  
QY 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110  
Db 309 ATGCAAGTAGAACTCGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 368  
QY 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGlyValThrGlyThrPro 130  
Db 369 GAGCTGACACTTGCATAATTTCTTTCAGGACACATGAAGAAATGACCCTGGGACACA 428  
QY 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150  
Db 429 ATTCATGTTTGTCCCGAACACAGATCAGAGAGCGGGGTGATTACAGTGAATGCTAAG 488  
QY 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170  
Db 489 GCCTACAGACCTTCACATGCGAGATGCAACTTATGACTTCAAAATACGGTGTAGAGCAGT 548  
QY 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190  
Db 549 CAGGAGGTGGAAGATCATCATCAGAGAGACCATTTGAAGGGTGGCTGCGAGGAGCTCT 608  
QY 191 ProLysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210  
Db 609 GCMAAGAAATTCCTAAGCTCAATCTGGAGTAGAGATCTTGGCGTGTGTGCTCAAGGTG 668  
QY 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230  
Db 669 CATCAAGTTGACTACCAAGAGATGCCCTTGAATATGACATCTGTAACCAATGGAACAGATA 728  
QY 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250  
Db 729 GAAAGCAACATGTTAGATGTTCTGATCCAGAAATATGCAGAGAGATGATTGATGACTCTC 788  
QY 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270

Db 789 GATAAAGTACGAGTTAGAGTTCGATTGGTGTGGTGCATGCATGCAATGCAAGAAAT 848  
Qy 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290  
Db 849 GTTCCTCGGAGATTGGCTCTCTGTATTGACAAACTTGAGGCTGAAATTGGCGAAGCT 908  
Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr 309  
Db 909 ATGCTTTCTCTCTCTGCAAGCAAGGGGTTTGAGATCGGCAGTGGGATTGTGTTCAC 965

## RESULT 4

US-09-743-207-5  
; Sequence 5, Application US/09743207  
; Patent No. 6653531  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/09/743,207  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/093,611  
; PRIOR FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 1015  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-743-207-5

Alignment Scores:  
Pred. No.: 1,48e-111 Length: 1015  
Score: 1066.00 Matches: 201  
Percent Similarity: 93.83% Conservative: 27  
Best Local Similarity: 82.72% Mismatches: 15  
Query Match: 48.45% Indels: 0  
DB: 4 Gaps: 0

US-10-677-179-8 (1-429) x US-09-743-207-5 (1-1015)

Qy 186 AlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSer 205  
Db 8 GTCGAGGAGCTGTTGCAAGAAAAATCTTAAGCTGAAATGCGAGTAGAGATTCTAGCA 67  
Qy 206 PheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerVal 225  
Db 68 TTGCTTCCAAAGTCATCAAGTGGTACTCTCTGAGAGCGAGTTGATTATGAAACTCTT 127  
Qy 226 ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys 245  
Db 128 ACCCTGGATCAGATAGAGAGCAACATTTCTAGATGCTCTGATCCAGATATGCACAGAAG 187  
Qy 246 MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThr 265  
Db 188 ATGATTGATGCAATTGATTAAGTACGAGTTAATGGGAATTCGATTGGTGGGTGTCACA 247  
Qy 266 CysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSer 285  
Db 248 TGCATTGCCAAGAAATGTTCTCTGTGGGCTTGCTCTCTGTTATTTGACAAACTGAGCT 307  
Qy 286 GluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGly 305  
Db 308 CTACTGGCAAGAGCTATGCTTCTCTCTGCAAGCAAGGGGTTTGAGATCGGTAGTGA 367  
Qy 306 PheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAla 325  
Db 368 TTTCAGGCTACTGACCTAATGGAAGTGAGCAATACGATGAGTTCTATATGACGAGGCT 427  
Qy 326 GlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyIleSerAsnVal 345

Db 428 GGAATGTAAGAACACCAATCGCTCGGGCGTGTACAGGGAGGATATCAATGGT 487  
Qy 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365  
Db 488 GAAACTATATACTTCAAAAGTAGCTTCAAGCCAAACAGCAACTATTGGGAAGAAACAA 547  
Qy 366 ThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCys 385  
Db 548 ACTGTAAACAAAGGATCATGAGGATATCGAACTTCTGACAAAGGGGTGCGCATGACCCATGT 607  
Qy 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405  
Db 608 GTCGTCCTCGGGCTGTTCCAATGCTGAGACGATGCTGCATTGGTCTCATGACACAG 667  
Qy 406 LeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluPro 425  
Db 668 CTGATGGCACATGTTGCTCAGTGGGAGATGTTCCCGCTGAACCTCGCCCTACAAGAACA 727  
Qy 426 ValGlySer 428  
Db 728 ATCGGCTCC 736

## RESULT 5

US-09-489-039A-5375  
; Sequence 5375, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5375  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5375

Alignment Scores:  
Pred. No.: 6.94e-88 Length: 1134  
Score: 859.00 Matches: 180  
Percent Similarity: 64.99% Conservative: 65  
Best Local Similarity: 47.75% Mismatches: 112  
Query Match: 39.05% Indels: 20  
DB: 4 Gaps: 6

US-10-677-179-8 (1-429) x US-09-489-039A-5375 (1-1134)

Qy 40 HisArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyr 59  
Db 16 CACGCCAACATAACGAAAAAGATTGGAGCCGTGATGGCAGGAATACATAATTGGACAATC 75  
Qy 60 PheGlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGly 79  
Db 76 TTTCGCGTCACTACCTTCGCGCAATCGCACGCTCTGGCGCTCGGCTGTCATCGTTGACGGC 135  
Qy 80 CysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArg 99  
Db 136 GTGCGCGCGGATCTCCGCTACCCAGCCACCTGACAGCAGCAGCTCGATCGTCTGCGC 195  
Qy 100 ProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSer 119  
Db 196 CCGGGTACTCTCGGTTACACCAACGCGCGGTGCGAGCGCGGATCAGGTCAAAATCTCTCC 255  
Qy 120 GlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAsp 139  
Db 256 GCGCTCTTCGAAGGCGTTTACCACCGGACGAGTATCGGTCTGCTGATTGAGAACACCGAT 315

```
QY 140 GlnIleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAla 159
Db 316 CAGCGTTCCGAGGATTACGCGCGCATCAAGATCTGTTCCGTCCGGCCATCGGATTAC 375
QY 160 ThrTyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArg 179
Db 376 ACCTACGAGCAAAAGATATGCTCTCGCGCATATCGCGCGCGCGTCTTCTTCTGCGCGC 435
QY 180 LysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysCys 199
Db 436 GAAACCGCGATGCGGTAGCGCGCGCGCGCGATTCGCAAAAATACCTTCGCGCGAAGTTC 495
QY 200 GlyLeuGluLeuLeuSerPheValSerLysValHisGlnValValLeuProGluAlaAspAla 219
Db 496 GGCATCGTCAATCGCGCTGCTCGCGCATCGCATGCGCGATATT-----CGCTGGCCATC 549
QY 220 ValAspTyrClySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAsp 239
Db 550 AAGACTGG-----GATCAGGTTGACGAGAACCCCTTCTTCTGTCGGAC 594
QY 240 ProGluTyrAlaGluLysMetIleAspAlaIleAspArg-----ValArgVal 255
Db 595 CCGGAT-----AAATCGACGCCCTCGATGAAGTATGTCGCGCGCTGAAAAA 642
QY 256 ArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeu 275
Db 643 GAGGCGACTCCATTCGCGCGAAGTACGGTGTGCTGACGCGCTACCGCGACGCTC 702
QY 276 GlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIlePro 295
Db 703 GGTGAGCCGCTCTTCACCGCTGATGCGCATATCGCCACGCGCTGATGAGCATCAAC 762
QY 296 AlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGlu 315
Db 763 GCGGTGAAGGGTGAATATGCGACGCGCTTCGAGGTGTGAAGTTCGCGCGACGCGAA 822
QY 316 HisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSer 335
Db 823 AACCGCGATGAG-----ATCACTAAGCGCGG-----TTCCAGACCAACACGCGC 867
QY 336 GlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysValAlaPheLys 355
Db 868 GCGCGCATCTTTCGCGCATCAGCAGCGCGCCAGCATGCTGCGCAACATCGCGCTGAAG 927
QY 356 ProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGlu 375
Db 928 CTTACCTCCAGTATTACCGTTCCCGTCAACATTAATTCGTTTGGCGAAGTGGAG 987
QY 376 LeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValProValValGlu 395
Db 988 ATGATCAACCAAGGGCGTCACGATCGTGTGCGCATCGCGCGCTACCGATCGCTGAG 1047
QY 396 SerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db 1048 GCGATCTCGCGATCTGTTCTGATGATCACTTTATCGTTCGCGCGCGCGAG 1098
```

## RESULT 6

```
US-09-543-681A-120
; Sequence 120, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 120
; LENGTH: 1089
; TYPE: DNA
```

```
; ORGANISM: Proteus mirabilis
US-09-543-681A-120
Alignment Scores:
Pred. No.: 6,91E-87 Length: 1089
Score: 850.00 Matches: 167
Percent Similarity: 66.02% Conservatives: 72
Best Local Similarity: 46.13% Mismatches: 111
Query Match: 38.64% Indels: 12
DB: Gaps: 3
```

US-10-677-179-8 (1-429) x US-09-543-681A-120 (1-1089)

```
QY 52 SerGlyAsnThrPheClyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 71
Db 7 GCGGGAACAGTATCGGCAATATTATTAGAGTAACCACTTTTGGTGAGTCTCATGGCACA 66
QY 72 GlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeu 91
Db 67 GCGTTAGGTTGCATTGTCGATGCTCTCTCCCGGACTACCTTTAACGCAAGCGGATCTA 126
QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
Db 127 CAAGTTGATTAGATAGACGTAACCCGGAACCTTCACGTTATACCAACACACAGGTAGAG 186
QY 112 ThrAspThrCysLysIleLeuSerGlyThrHisGlyValThrThrGlyThrProIle 131
Db 187 CCGTATCAAGTCGGTATTTATCGGGTGTATTTAATGGTGAACACGGAACCAAGTATT 246
QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
Db 247 GGAATTATTAAAGAAATACGGATCAGCGCTCTCAAGATTATAGCGAATTAAGATGTA 306
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
Db 307 TTCCGCGCCAGGCGATCGACACTACACCTATGAACAGAAATATGGTTTACGTGATTATCGT 366
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyValAlaLeuPro 191
Db 367 GCGCGGAGCGCTCTCTGTCGAGAAACCGCATCGCTGTCGACGCGGTCTATTGCC 426
QY 192 LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db 427 AAAAATAACCTTAAACAAAAATTTGTTATGTAAGTAAAGGTTACTTATCTCAATTAGGG 486
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
Db 487 CCGATTAGTTGT-----GAGTTAGTTGATGTCATTTGTA-----GAA 525
QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
Db 526 ACCAACCCATCTTTTTCGCCAGATCCTTCTCGCTAGATGCCCTTGATGAATATATATGCA 585
QY 252 ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaAlaArgAsnVal 271
Db 586 GCGCTAAAAAAGAGGGTAATTTCTATAGTGCCAAAGTTACCGTGTGTCGAGAGGTGTA 645
QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
Db 646 CCGCAGGATTAGGTGAACCCGCTTTTTCATAGACTTGTATGCGGATTTGGCTCATGCGTTA 705
QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
Db 706 ATGAGCATCAACAGTAGTAAAGCGCTTGAAATTCGAGATGGTTTTTGATGTGGTGCACCTTA 765
QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
Db 766 AAAAGGACTGAAACCGCGATGAAATTAACCAAGAGGATTTAGCAGT-----813
QY 332 ThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLys 351
Db 814 ---AATCAGCGGGGGCGGTATTAGCGGCATTAGTAGTGGTCAACCTTATTATTATTCACAC 870
```





Db 1159 ACGCAACGGCGCAA 1174

## RESULT 8

US-09-064-693A-26  
; Sequence 26, Application US/09064693A  
; Patent No. 6210937  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Thomas E.  
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY  
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION  
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: W. Gary Goodson  
; ADDRESSEE: INEEL--Lockheed Martin Idaho  
; ADDRESSEE: Technologies Co.  
; STREET: P.O. Box 1625  
; CITY: Idaho Falls  
; STATE: Idaho  
; COUNTRY: USA  
; ZIP: 83415-3810  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
; COMPUTER: Toshiba Satellite Pro T2150CDS  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: Word Perfect 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/064,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: W. Gary Goodson  
; REGISTRATION NUMBER: 22,387  
; REFERENCE/DOCKET NUMBER: LIT-PI-296  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (208)526-9469  
; TELEFAX: (208)526-8339  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-09-064-693A-26

Alignment Scores:  
Pred. No.: 1,46e-85 Length: 4530  
Score: 847.00 Matches: 178  
Percent Similarity: 65.30% Conservative: 61  
Best Local Similarity: 48.63% Mismatches: 106  
Query Match: 38.50% Indels: 21  
DB: 3 Gaps: 5

US-10-677-179-8 (1-429) x US-09-064-693A-26 (1-4530)

QY 52 SerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyxGlyGluSerHisGlyGly 71  
Db 127 GCTTGAACACCAATTCGACCACTCTTCGCGTAACCACTTCGCGGAATCGCAGCGGTG 186  
QY 72 GlyValGlyCysValleserGlyCysProProArgileProleuthrGluAlaSerLeu 91  
Db 187 GCGCTCGGTGTCATCGTCGATGCTGTCGCGGAGGCAATTCGCTGACGGAAGCGGACCTG 246  
QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgileThrSerThrArgLysGlu 111  
Db 247 CAACATGACCTGACCGCTGCTGCGCTGGGACATCGCGCTATACCAACCGGCGCGGAG 306  
QY 112 ThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProle 131  
Db 307 CCGGATCAGGTCAAATCTCTCCGGTGTGTTTGAAGGCGGTACTACCGGCAACGACATT 366

QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaVal 151  
Db 367 GCGTGTGATCGAAACACTGACCGCTCTCAGGATTACAGTCGCGATTAGGACGT 426  
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171  
Db 427 TTCCTGCCAGGCCATCGCGATTACACCTACGAAACAAATAATCGGTCTGCGCGATTATCGC 486  
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191  
Db 487 GCGCGTGGACGTTCTTCCCGCGGAAACCGCATCGCGCTGCGCGCAGGAGCTATTGCC 546  
QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211  
Db 547 AAAAATAATCTCGCGAGAAATTTGTTATTTGAATCCGTGGCTGGCTGACCCAGATGGGC 606  
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231  
Db 607 GACATT-----CGCTGATATCAAGACTGG-----TCGCAAGTTCGAG 645  
QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaLeu 251  
Db 646 CAAAATCCGTTTTTTTTCGCGGACCGCGAC-----AAAATCGACCGCTTAGAC 693  
QY 252 Arg-----ValArgValArgGlyAspSerValGlyValIleThrCysVal 267  
Db 694 GAGTTGATCGCTGCGCTGAAAAGAGCGGCGACTCCATCGGCGCTAAAGTCAACCTGTTGT 753  
QY 268 AlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeu 287  
Db 754 GCCAGTGGCTTCCTCGCGGACTTCGCGAGCGGTCTTTGACCGCTGATCGACATC 813  
QY 288 AlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAla 307  
Db 814 GCCCATGGCTGATGAGCATCAACCGCGGTGAAAGCGGTGAAATTTGGCGACGGCTTTGAC 873  
QY 308 GlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySer 327  
Db 874 GTGTGGCGCTGCGCGGAGCGCAGCAACCGCGATGAATCACCAGAAC-----921  
QY 328 ValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIle 347  
Db 922 ---GGTTTCCAGAGCAACCATCGCGGCGGCATTCTCGCGGTATCAGCAGCGGCGACAA 978  
QY 348 ValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrVal 367  
Db 979 ATCATTGCCCATATGCGCGCTGAAACCGACCTCCAGCATTTACGCTCGCGGTGCTACCAT 1038  
QY 368 SerArgGluArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAla 387  
Db 1039 AACCGCTTTGGCGAAGAAGTTGAGATGATACCAAGGCGCGTCACGATCCCTGTGTGGG 1098  
QY 388 ProArgAlaValProValValGluSer-MetAlaAlaLeuValLeuMetAspGlnLeuMe 407  
Db 1039 ATCCGCGCAGTGGCGATCGCAGAGCGAATGCTGCGGATCGTTTAAATGGATCACCCTGT 1158  
QY 407 AlaHisValAlaGln 412  
Db 1159 ACGCAACGGCGCAA 1174

## RESULT 9

US-09-790-988-1/c  
; Sequence 1, Application US/09790988  
; Patent No. 6632935  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEMI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988



```
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Alignment Scores:
Pred. No.: 6,5e-82 Length: 640681
Score: 845.00 Matches: 172
Percent Similarity: 64.84% Conservatives: 64
Best Local Similarity: 47.25% Mismatches: 116
Query Match: 38.41% Indels: 12
DB: 4 Gaps: 3

US-10-677-179-8 (1-429) x US-09-790-988-1 (1-640681)

Qy 50 LysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHis 69
Db 102991 AAATGCTCTGGAATACAAATTCGGGAAATATTCGTGTAAACCACTTTTGGCGAGTCACAC 102932

Qy 70 GlyGlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAla 89
Db 102931 GGAGAAAGCAATAGGGTGCATATTAACGGAACACCTCTCGTCTTGAATATTCCTGTAAA 102872

Qy 90 AspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArg 109
Db 102871 GATTGCAATATGATTGAATCGTAGAAGACCAAGTACTTCCGTTACACGACTTTGGCG 102812

Qy 110 LysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThr 129
Db 102811 CGTGAACCTGATGAAGTAAATATCTTCCGGTATATTAACGGGGTCACAACTGGTACT 102752

Qy 130 ProLeuValIleValProLeuThrAspGlnIleGlySerAspHisArgGluIleAla 149
Db 102751 AGTATGGTTTAAATATATTAATCATGATCATAGCTCAAGACCTATAGCGATATAAAG 102692

Qy 150 AsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAla 169
Db 102691 AATTATTTCGACCGGACATCCGATTAATCTATGAAAAAATAATATGGAATTAGAGAT 102632

Qy 170 ValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAla 189
Db 102631 TATCGCGGGGAGGTAGATCTTCTGCTCGGAAACTGCTATGAGAGTTGACGACGAGCT 102572

Qy 190 LeuProLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLys 209
Db 102571 ATAGCAAAAAATACCTTAATAGAAATATGGGATAACTATTTCGACATATTGTTCAGCA 102512

Qy 210 ValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGln 229
Db 102511 ATGGGCAATACAAATGCCCT-----TTTAAATCTTGGCAAGAA 102473

Qy 230 IleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAla 249
Db 102472 GTGCAAAATAATCTTTTCTGTTCTGCTCGTGAATAATTTTATGCACTAGAAAAATTTA 102413

Qy 250 IleAspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArg 269
Db 102412 ATTAATATCTCAAAAAAATAGTGTATTCGATTCGTCGTAATATACAAATTAATCTGTA 102353

Qy 270 AsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLys 289
Db 102352 AATATCCCTGTAGGACTTGGAGAACCCAGTTTTTGTATCGTCTGATCGTCTGATCTATCACAC 102293

Qy 290 AlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr 309
Db 102292 GCCTTGATGATTAATTAATCCGCAAAAGAGGATAGAAATTTGGAGACGGTTTTTTCAGTAATA 102233

Qy 310 AspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArg 329
Db 102232 AATCAACGAGGAGTGAACATCTGTATGAA-----ATTACGCCACAGGATTTTAA--- 102182

Qy 330 ThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHis 349
Db 102181 -----ACTAATCATCTCTGGTGGTATTCTAGTCTGATTAATGTCGAGAAATTTGTA 102128

Qy 350 PheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArg 369
Db 102127 CTAAAGTAGCATTTAAACCTACATCAAGTATTTCGAAAGCAGGCAATACAAATAATA 102068

Qy 370 GluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArg 389
Db 102067 AATTAACGAAAGTTCAAATAGTTACTTAAGGACGACATGATCCATGTGTAGTTTACGC 102008

Qy 390 AlaValProValValGluSerMetAlaLeuValLeuMetAspGlnLeuMetAlaHis 409
Db 102007 GCTGTTCCATCACTGAAGCAATGTTAGCAATTTGATTAATGGATCATTTATTAAAGATT 101948

Qy 410 ValAlaGlnCys 413
Db 101947 CGCGCAATATGC 101936

RESULT 10
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PBI86P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.: 1.12e-80 Length: 1830121
Score: 840.50 Matches: 166
Percent Similarity: 64.74% Conservatives: 69
```

Best Local Similarity: 45.73% Mismatches: 121  
Query Match: 38.20% Indels: 7  
DB: Gaps: 3

US-10-677-179-8 (1-429) x US-09-557-884-1 (1-1830121)

```
QY 52 SerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyGlyGluSerHisGlyGly 71
DB 211427 GCTGGTAATACAACTTTCCTGGTGACAACTTTGGAGAGTCACATGGTATT 211486

QY 72 GlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaLeu 91
DB 211487 GCATAGGCTGATTCCTGATGCGCGCCACCAATTCGGAATTTATCCGAGAAATATT 211546

QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
DB 211547 CACCGAGATTAGATCGTGTAAACGAGGAACATCTGATATACGACGCTCGTGTGAA 211606

QY 112 ThrAspThrCysAllysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIle 131
DB 211607 GATGACGAAGTTCAAAATTTTATCTGCTGTGTTTGAAGGAAAAACACACAGCACAAGTATT 211666

QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaVal 151
DB 211667 GCGATGATCAATTAATAAATGGAGATCAGCGTTCCAGAGATTATGTCACATTAAGATCGT 211726

QY 152 TyrArgProSerHisAlaAspAlaThrTyArgPheLysTyGlyValArgAlaValGln 171
DB 211727 TTCGCGCCAGGTCATCGGATTTTACCTATCAGCAAAAGTATGGAATCGGTATATCGT 211786

QY 172 GlyGlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
DB 211787 GCGGTGGCGGTTCGTACACGCTGAAACAGCGATCGGTGTCGAGGGCTATTGGC 211846

QY 192 LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
DB 211847 AAAAAATATTACCGCAACATTTTGGCATTTAGTTCGAGGTTTTTTTAAAGCAAAATCGT 211906

QY 212 GlnValValLeuProGluAspAlaValAspTyGlySerValThrLeuGluIleGlu 231
DB 211907 AATATAAAATTTGCTCCGCGAGAAAGTG-----GGACAAATGATTTGGGAAAGGTAAAC 211960

QY 232 SerAsnIleValArgCysProAspProGluTyAlaGluLysMetIleAspAlaIleAsp 251
DB 211961 AGTAATCCATTCTTTGTCTCTGATGAAGTGGGTAGAAAAATTCGATGAATGATCGT 212020

QY 252 ArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnVal 271
DB 212021 GAACTTAAAAAAGAGGAGATTTCTATTGGCGCAAACTTACTGTTATTTCAGAAAAATGTA 212080

QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
DB 212081 CCTGTAGGATTTGGCGAGCGAGTATTGTACCGTTTAGATGCGCGATCTTCTCCACCGATTA 212140

QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
DB 212141 ATGGGAATTAATGCATGAAGGTGTAGAAATTTGGCGATGGCTTGTCTGTGGTGAACA 212200

QY 312 ThrGlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgThrArg 331
DB 212201 CGAGGTTCCGAAACATCGTGATGAA-----ATGACACCTTAATGGC-----TTTGAA 212245

QY 332 ThrAsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPheLys 351
DB 212246 AGTAATCATCGCGCGGTATTATTAGCGGAATTTAGTTCAGGACCAACCAATTAATCCCACT 212305

QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
DB 212306 ATTGACCTAAACCAACTCAAGCATTTACGATTCGTTGCTGCTGATCAATCTTAATGGT 212365

QY 372 GlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal 391
DB 212366 GAAGCGGTAGAAATTGTAACAAAAGTCTGCACGATCTCTTGTGTGGGATTCGTGCTGTG 212425
```

```
QY 392 ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
DB 212426 CCAATTGGGAGCTATGTCGCGATTGCTTTATTAGATCATCTCTTACGTTTAAAGCA 212485

QY 412 GlnCysGlu 414
DB 212486 CAGTGTAAA 212494

RESULT 11
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PBI86P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
```

Alignment Scores:

Pred. No.:	1,12e-80	Length:	1830121
Score:	840.50	Matches:	166
Percent Similarity:	64.74%	Conservative:	69
Best Local Similarity:	45.73%	Mismatches:	121
Query Match:	38.20%	Indels:	7
DB:	4	Gaps:	3

US-10-677-179-8 (1-429) x US-09-643-990A-1 (1-1830121)

```
QY 52 SerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyGlyGluSerHisGlyGly 71
DB 211427 GCTGGTAATACAACTTTCGTTGGACAACTTTTCGTTGGAGAGTCACATGGTATT 211486
```

Qy	72	GlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeu	91
Db	211487	GCATTAGGCTGTATCGTTGATGGCGTCCCAAAATCTCGAATTTATCCAGAAACATATT	211546
Qy	92	GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu	111
Db	211547	CAGCAGCATTTAGATCGTCGTAAACACAGAACATCTCGATATACAGACCCCTCGTCGTGAA	211606
Qy	112	ThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIle	131
Db	211607	GATGACGAGTTCAAATTTTATCTGCTGTGTTGAAGAAACACACAGGCACAGTATT	211666
Qy	132	LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal	151
Db	211667	GGGATGATCAATAAAAATGGAGATCAGCGTTCCGAAGATTATGGTGACATTAAGATCGT	211726
Qy	152	TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln	171
Db	211727	TTCCGCCAGGTCATCGCGATTTTACCTATCAGCAAAAGTATGGAATCCGTGATTATCGT	211786
Qy	172	GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro	191
Db	211787	GCCGTTGGCGTTCGTACGACGTGAACAGCGATCGCGTTGCTGCAGGGGCTATTGGG	211846
Qy	192	LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis	211
Db	211847	AAAAAATATTTACGCCAAACATTTTCGCATTCAGGTGCGAGGTTTTTTAAGCCAAATCGT	211906
Qy	212	GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu	231
Db	211907	AATATAAAATTTGCTCCGACAGAAAGTG-----GGACAAATTTGATGGGAAAGGTAAAC	211960
Qy	232	SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp	251
Db	211961	AGTAATCCATCTTTTGTCTCTGATCAAAAGTCGGGTAGAAAATTCGATGAATGATCCGT	212020
Qy	252	ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnVal	271
Db	212021	GAACCTAAAAAAGAGAGATTTCTATTGGCGCAAACTTTACTGTTATTTCAGAAATGTA	212080
Qy	272	ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet	291
Db	212081	CCTGTAGATTGGCGCGCAGCATTTTGACGTTTAGATGCCGATCTTGCTCAGCATTA	212140
Qy	292	LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu	311
Db	212141	ATGGCAATTAATGTCAGTAAAGGTGTAGAAATTTGCGATGGCTTTGCTGTGTGAACAA	212200
Qy	312	ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg	331
Db	212201	CGAGGTTCGGAAACATCGTGATGAA-----ATGCACCTTAATGGC-----TTTGAA	212245
Qy	332	ThrAsnArgSerGlyGlyValGlnGlnGlyIleIleSerAsnValGluIleValHisPheLys	351
Db	212246	AGTAATCATGCGAGCGGTATTTTAGCGGAATTAGTTTCAGGACCAACCAATTCGCCACT	212305
Qy	352	ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg	371
Db	212306	ATTGCATTAACCAACCACTTCAAGCATTAACGATTACGATTCCTGGTGGTTCAATCAATCT	212365
Qy	372	GlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal	391
Db	212366	GAAACCGTAGAAGTTGTAAACAAAGGTGCTCAGATCCTCTGTGTGGGGATTCGTGCTGG	212425
Qy	392	ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla	411
Db	212426	CCAATTGCCGAGCTATGTTGGCGATTGCTTATTAGATCATCTCTACGTTTTTAAGGCA	212485
Qy	412	GlnCysGlu	414
Db	212486	CAGTGTAAA	212494

```
Db      529 AATGTCGTGAAGAAATGGTTAAAGTGATTCAAAAATATCGTCATGCTAAAGATTCATT 588
QY      261 GlyGlyValIleThrCysValAlaAraAsnValProArgGlyLeuGlySerProValPhe 280
Db      589 GGTGGTGTCTGCTACTGTGTATCAGAAACTGTCCTAATGGATAGGTGAACATGTTTC 648
QY      281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db      649 GATAAATTGGAAGCTAAATTTGGCTCATGCTCATGTTATCATTTACCCGCTACCAAGGGTTT 708
QY      301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
Db      709 GAATTTGGTTCGTGATTGAAGGTATCAAGATCCCTGGTTCAAAACATAACGATGCATTC 768
QY      321 TyrMetAspLys---AlaGlySerValAlaThrArgThrAsnArgSerGlyGlyValGln 339
Db      769 TATTATGATGAAATTTCCGAGAGATTAAGAACTGAAACCAACAATAGTGTGTATCCAA 828
QY      340 GlyGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSer 359
Db      829 GGAGGAATTTCAACCGTGAAACATTTATTTCTCAGTTGCTTCAAAATCGGCTGCTACT 888
QY      360 IleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuAlaArg 379
Db      889 ATCAGTCAAGAACCAAGAACTGCTACTTACGATGTTAAAGTGGTGTGCTGCTAGA 948
QY      380 GlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAla 399
Db      949 GGTAGACACATCCCAAGTGTACCCCAAGACGAGTTCCTCAATTTGTAAGCCATGACTGCA 1008
QY      400 LeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db      1009 TTAGTATTATGTACGAATATATGATTCACAAGCTAGA 1047

RESULT 13
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ 'ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103398)..(103398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
```

```

; NAME/KEY: misc feature
; LOCATION: (67435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (71352)..(71352)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (85539)..(85539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

```

Alignment Scores:
Pred. No.: 3.5e-76
Score: 800.50
Percent Similarity: 61.52%
Best Local Similarity: 44.76%
Length: 1664976
Matches: 171
Conservative: 64
Mismatch: 122

```

```

Query Match: 36.39% Indels: 25
DB: 4 Gaps: 6
US-10-677-179-8 (1-429) x US-08-916-421B-1 (1-1664976)
Qy 47 LeuGluValLysAlaSerGly-----AanThrPheGlyAen 58
Db 357409 TTAATGTAAACTAAGCAATTAATAATTTTGGTGACATTAATGAACACCTATGGGGAT 357468
Qy 59 TyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSer 78
Db 357469 ATGTTTAGAGTTACAGTTTGGAGAAAGTCATGGAAGGCTGTTGGAGCAGTGTGTCAT 357528
Qy 79 GlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArg 98
Db 357529 GGATGTCAGCTAATCTGCCTTTATCTGAAGAGATATCAAAAAGAGCTTGACAGAGA 357588
Qy 99 ArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeu 118
Db 357589 AGACGAGGCGAGCATCTTCTCAACCAAGAAAGAGAGGATAAGTTGAATCTTA 357648
Qy 119 SerGlyThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThr 138
Db 357649 TCAGGAATTTTGGAGGGAACACTACTGGAGCTCTATTTGCTCAATAGTCTATAACAAA 357708
Qy 139 AspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAsp 158
Db 357709 AACATGAGACCTAAGATTTACTCAAAATTAAGAGATACACCAAGACTGGACATGCAGAT 357768
Qy 159 AlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGly 178
Db 357769 TTAACCTATAGATTGAAGTATAAAACTATGATTATAGGGGAGGAGGAGGCAAGTGT 357828
Qy 179 ArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLys 198
Db 357829 AGAGTAACGATAGGCGCATGTTATTGGAGGAGCTATTGCTAAAAAGCTTCTATCTTACACA 357888
Qy 199 CysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal-----213
Db 357889 TACAACATAAAATTTATTTGGTTATACCAATAAAGATTGGAAGATTGAAGGAGATTCAGC 357948
Qy 214 -----ValLeuProGluAspAlaValAspTyrGlySerVal-----ThrLeuGluGln 229
Db 357949 TACTATAAAATCCAGAGGTTTTTGAATAATCAAAATCCCTAGAGAGATTAAATAGAGATT 358008
Qy 230 IleGluSerAenIleValArgCysPro-----AspProGluTyrAlaGluLysMet 246
Db 358009 ATTGAAAGTAAATCCATTGAGATGTCCATCAATGAATGAGAAAGAGATGGAGGATGTGT 358068
Qy 247 IleAspAlaIleAspArgValArgGlyAspSerValGlyGlyValIleThrCys 266
Db 358069 TTAAGGCAATGGAA-----AATAAAGATAGTGTGGAGGAGTTGTTGAAATT 358116
Qy 267 ValAlaArgAenValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu 286
Db 358117 GTTGCATTAAATGTTCTCTGTTGGAGTTGGAATCCAAATATTCAATAAGTTAAATGGAGAA 358176
Qy 287 LeuAlaLysAlaMetLeuSerIleProAlaSerAenGlyPheGluIleGlySerGlyPhe 306
Db 358177 TTGGCAAGAGCTTTAATGAGTATAAATGCTGTTAAAGAGAGTTGAGATAGGGCGCTGTTT 358236
Qy 307 AlaGlyThrAspLeuThrGlySerGluHiAenAspGluPheTyrMetAspLysAlaGly 326
Db 358237 AAAGCGCTGAGATGTATGGAAGTGAGATGAACCATGAGATGTTATTTTGTAGACACAAA 358296
Qy 327 SerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAenValGlu 346
Db 358297 AATATAAGATTCAAAACAAACAACTGCGGTGGCATATTGGGAGGAGATTAGTGTGGAAT 358356
Qy 347 IleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGluAsnThr 366
Db 358357 CCAATAGTTTAAAGATTGCAAGTAAAGCCAACTTCAATAGGTAAAGCAAAAAACC 358416

```

QY 367 valSer---ArgGluArgGlnAenValGluLeuLeuAlaArgGlyArgHisaspProCys 385  
Db 358417 ATAAATTAAACCTTAGAAATGTTGAAATTTGAAGGAGACACGACCCAGTT 358476  
QY 386 valAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405  
Db 358477 ATAGTTCCAAGGATTGTTCCAGTGGCTGAAGCAATGCTGTATATACCTTAGCTGATTTG 358536  
QY 406 LeuMet 407  
Db 358537 ATGATT 358542  
RESULT 14  
US-09-692-570-1  
; Sequence 1, Application US/09692570  
; Patent No. 6797466  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
; TITLE OF INVENTION: jannaschii  
; FILE REFERENCE: PB275C1  
; CURRENT APPLICATION NUMBER: US/09/692,570  
; CURRENT FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; PRIOR APPLICATION NUMBER: US 08/916,421  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (103998)..(103998)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (148948)..(148948)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (163385)..(163385)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (191989)..(191989)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (191995)..(191995)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (231980)..(231980)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234187)..(234187)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234220)..(234220)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (234814)..(234814)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (309398)..(309398)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (309418)..(309418)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (312837)..(312837)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (312993)..(312993)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (319226)..(319226)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (559167)..(559167)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (559241)..(559241)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (600992)..(600992)  
; OTHER INFORMATION: n equals a, t, c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (622708)..(622708)  
; OTHER INFORMATION: n equals a, t, c, or g

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:
Pred. No.: 3.5e-76
Score: 800.50
Percent Similarity: 61.52%
Length: 1664976
Matches: 171
Conservative: 64

```

---

```

Best Local Similarity: 44.76% Mismatches: 122
Query Match: 36.39% Indels: 25
DB: 4 Gaps: 6
US-10-677-179-8 (1-429) x US-09-692-570-1 (1-1664976)
Qy 47 LeuGluValLysAlaSerGly-----AenThrPheGlyAsn 58
Db 357409 TTAATGTAAACTAAGCAATTTTGGTGACATTATGAACACCTATGGGGAT 357468
Qy 59 TyrPheGlnValAlaThrTyGlyGlySerHisGlyGlyValGlyCysValIleSer 78
Db 357469 ATGTTTAGAGTTACAGTTTTCGAGAAAGTCATCGAAAGCGTGTTCGAGCAGTTGTCAT 357528
Qy 79 GlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArg 98
Db 357529 GGATGTCAGCTAATCTGCCTTTATCTGAAGAGATATCCAAAAGAGCTTGACAGAGA 357588
Qy 99 ArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeu 118
Db 357589 AGACGAGGCGAGCATCTTCTCAACCAAGAAAGAGAGGATAAGTTGAANTCTTA 357648
Qy 119 SerGlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThr 138
Db 357649 TCAGGAATTTTGGAGGGAAAACTACTGGAGCTCTATTGCTCAATAGTCTATAACAAA 357708
Qy 139 AspGlnIleGlySerAspHisArgGluIleAlaAsnValTyArgProSerHisAlaAsp 158
Db 357709 AACATGAGACCTAAAGATTACTCAAAAATTAAGATACACCAAGACCTGGACATGCAGAT 357768
Qy 159 AlaThrTyraPhePheLysTyrglyValArgAlaValGlnGlyGlyArgSerSerGly 178
Db 357769 TTAACCTATAGATTGAAGTATAAAACTATGATTATAGGGGAGGAGGAGGCAAGTGGT 357828
Qy 179 ArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLys 198
Db 357829 AGAGTAACGATAGGCGCATGTTATTGGAGGAGCTATTGCTAAAAAGCTTCTATCTACACA 357888
Qy 199 CysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal----- 213
Db 357889 TACAACATAAAAATTTATTGGTTATACCATAAAGATTGGAAGAGTTGAAGGAGATTCAGC 357948
Qy 214 -----ValLeuProGluAspAlaValAspTyrglySerVal-----ThrLeuGluGln 229
Db 357949 TACTATAAAAATCCAGAGGTTTTTGAAAATCAAAAATCCTTAGAGAGATTAATAGAGATT 358008
Qy 230 IleGluSerAenIleValArgCysPro-----AspProGluTyraAlaGluLysMet 246
Db 358009 ATTGAAGTAAATCCATTGAGATGTCATCAATGAATGAGAAAGAGATGGAGGAGTATGTT 358068
Qy 247 IleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCys 266
Db 358069 TTAAAGGCAATGGAA-----AATAAGATAGTGTGGAGGAGTGTGTAATTT 358116
Qy 267 ValAlaArgAenValProArgGlyLysSerProValPheAspLysLeuGluSerGlu 286
Db 358117 GTTCATTAAATGTTCTGTTGGAGTTGGAATCCAAATATCAATAATGTTAAATGGAGAA 358176
Qy 287 LeuAlaLysAlaMetLeuSerIleProAlaSerAenGlyPheGluIleGlySerGlyPhe 306
Db 358177 TTGGCAAGAGCTTTAATGAGTATAAATCTCTTAAAGAGTGTGAGATAGGGCGTGTGTTT 358236
Qy 307 AlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyzMetAspLysAlaGly 326
Db 358237 AAAGCGGCTGAGTGTATGGAAAGTGAGATGAACCATGAGATGTTTGTATGAGACACAAA 358296
Qy 327 SerValArgThrArgThrAsnArgSerGlyGlyValGlnGlnGlyIleSerAsnValGlu 346
Db 358297 AATATAGATTCAAAACAAACAACTCGCGGTGGCATATTGGGAGGAAATTAGCTGTGAACT 358356
Qy 347 IleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThr 366
Db 358357 CCAATAGTTTTTAAGAAATTGCAGTAAGCCAAACACCTTCAATAGGTAAAGCAAAAAACC 358416

```



Qy	179	ArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLys	198
Db	433	CGCAGACCGGCATGCGCGTGGCGCGGGGCTATCGCAAGAATACTCG	489
Qy	199	CysGlyLeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAap	218
Db	490	CTGGGCATCCAGGTGCGGGCTACATGAGCCAGCTCGGGCCGATCGATTCCGTTCCAG	549
Qy	219	AlaValAspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysPro	238
Db	550	AGCTGGGAC-----AGCGTCGAGCAGAATGCCCTTCTTCAGCGCC	588
Qy	239	AspProGluTyrAlaGluLysMetIleAspAlaIleAspArgValAlaArgValArgGlyAap	258
Db	589	GACCCGACAGGTGCGGAGCTGAGGCCTACATGACCACTTGCCTGGCGGACACAGGAT	648
Qy	259	SerValGlyGlyValIleThrCysValAlaAlaArgAsnValProArgGlyLeuGlySerPro	278
Db	649	TCGGTCGGGGCAAGATCACCGTGTTGCCGAAGCGTGCCGCGGGCTCGGGCGAGCGG	708
Qy	279	ValPheAspLysLeuGluSerGluLeuValAlaLysAlaMetLeuSerIleProAlaSerAsn	298
Db	709	ATCTTCACCGCTGGACCGCGAATCGCGCATCGCTGTAGAGATCAACGCGGTGAAG	768
Qy	299	GlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAap	318
Db	769	GGCGTGAGATCGCGCGCGCTTCGCCAGCATCGCCAGCGCGCACCGAGCACCGCGAC	828
Qy	319	GluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyVal	338
Db	829	GAA-----CTGACCCCGCAAGGCTTCCTG-----TCGACAAATCGCGCGGCATC	873
Qy	339	GlnGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrPro	358
Db	874	CTCGCGGGATCTCTCGGCAGCCCATCGTCCGCCACCTGGCGGCTCGAACCGCACTCC	933
Qy	359	SerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuMetAla	378
Db	934	AGCATCACCATCCCGGGCGCTCGATCGATACCGCGCGGACGCGGTGGACATGATCACC	993
Qy	379	ArgGlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAla	398
Db	994	AAGGGCGCTCAGACCCGTCGGTCCGCATCGCGGCCACGCGCATCGCGAGCGGATGATG	1053
Qy	399	AlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln	412
Db	1054	GCATCTGCTGTCGACCAATTGCTGGCGCAGCGTGGGAG	1095

Search completed: August 27, 2005, 07:51:07  
Job time : 2473 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 27, 2005, 04:44:29 ; Search time 727 Seconds  
(without alignments)  
3861.053 Million cell updates/sec

Title: US-10-677-179-8  
Perfect score: 2200  
Sequence: 1 MTTPKPOQVAHSRARLAPR.....VAQCEMFALTALQBPVGSF 429

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-MODEL=frame+p2n.model  
-Q=/cgn2\_1/USPTO\_spool\_P/US10677179/runat\_25082005\_150241\_24891/app\_query.fasta\_1.583  
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10677179 @CGN 1 1 723 @runat\_25082005\_150241\_24891  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US10J\_PUBCOMB.seq:  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	1626	18	US-10-677-179-7
2	2090	95.0	1845	20	US-10-425-115-13837
3	2016.5	91.7	1496	18	US-10-425-114-30604
4	2011.5	91.4	1541	20	US-10-425-115-13835
5	1800.5	81.8	1792	20	US-10-425-115-171490
6	1797.5	81.7	1635	18	US-10-677-179-1
7	1790.5	81.4	1603	18	US-10-425-114-32244
8	1724	78.4	1353	19	US-10-437-963-99207
9	1512	68.7	1674	18	US-10-425-114-8982
10	1512	68.7	2094	18	US-10-424-599-1811
11	1454	66.1	1311	19	US-10-696-616-27
12	1257	57.1	966	18	US-10-677-179-11
13	1191.5	54.2	1089	17	US-10-369-493-26364
14	1186.5	53.9	1086	17	US-10-369-493-42555
15	1182.5	53.8	1086	17	US-10-369-493-43897
16	1066	48.5	1015	18	US-10-677-179-5
17	1035	47.0	1080	17	US-10-369-493-44740
18	981	44.6	1068	17	US-10-369-493-34106
19	938	42.6	1131	17	US-10-369-493-45609
20	934	42.5	1137	21	US-10-741-849-6185
21	898	40.8	1292	18	US-10-424-599-1812
22	888.5	40.4	1074	17	US-10-282-122A-12647
23	869.5	39.5	1421	17	US-10-369-493-27659
24	863	39.2	724	20	US-10-425-115-171487
25	862	39.2	1086	17	US-10-282-122A-42063
26	860	39.1	1086	9	US-09-815-242-6135
27	860	39.1	1086	17	US-10-369-493-47170
28	860	39.1	1086	17	US-10-282-122A-20439
29	860	39.1	3100	21	US-10-489-273-5
30	858	39.0	1056	17	US-10-282-122A-24843
31	857	39.0	1086	9	US-09-815-242-7611
32	855	38.9	1083	17	US-10-282-122A-23613
33	850	38.6	1083	17	US-10-282-122A-32881
34	846.5	38.5	1074	17	US-10-282-122A-30794
35	845	38.4	640681	9	US-09-790-988-1
36	843	38.3	1080	17	US-10-369-493-44842
37	840.5	38.2	1071	14	US-10-260-877-127
38	840.5	38.2	1074	9	US-09-815-242-6884
39	840.5	38.2	1074	17	US-10-282-122A-21915
40	840.5	38.2	1077	17	US-10-369-493-44300
41	840.5	38.2	1830121	17	US-10-329-670-1
42	840.5	38.2	1830121	20	US-10-158-865-1
43	840.5	38.2	1830121	22	US-10-981-687-1
44	840	38.2	1086	17	US-10-282-122A-41210
45	830.5	37.8	726	19	US-10-767-701-9032

# ALIGNMENTS

RESULT 1  
US-10-677-179-7  
; Sequence 7, Application US/10677179  
; Publication No. US20040082050A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Pember, Stephen O.  
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
; FILE REFERENCE: BB-1159-A  
; CURRENT APPLICATION NUMBER: US/10/677,179  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US/09/743,207  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: PCT/US99/16353  
; PRIOR FILING DATE: 1999-07-20

! PRIOR APPLICATION NUMBER: 60/093,611  
! PRIOR FILING DATE: 1998-07-21  
! NUMBER OF SEQ ID NOS: 14  
! SOFTWARE: Microsoft Office 97  
! SEQ ID NO 7  
! LENGTH: 1626  
! TYPE: DNA  
! ORGANISM: Zea mays  
US-10-677-179-7

Alignment Scores:  
Pred. No.: 2,91e-239 Length: 1626  
Score: 2200.00 Matches: 429  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-677-179-8 (1-429) x US-10-677-179-7 (1-1626)

Qy 1 MetThrThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20  
Db 79 ATGACGACGGTCCCAAGCCACAGCAGGTGGCGCACTCACGGGCACGGCTCGCACCCCGC 138  
Qy 21 AlaileGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40  
Db 139 GCGATCGGGCTTGTCTGGAGTTTGCCCCAGCCCTCTCTCCCTCCGCTTCGGCGTGAC 198  
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPhe 60  
Db 199 CCGTCCGCCACTGCTCGCGCTAGAGGTGAAGCATCTGGAACACAGTTTGGAAACTACTTT 258  
Qy 61 GlnValAlaThrTyrGlySerHisGlyGlyGlyValGlyCysValIleSerGlyCys 80  
Db 259 CAGGTTCGAACCTATGGTGAATCTCATGGGGTGGTGTGGTGTATCAGTGGTTGT 318  
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100  
Db 319 CCACCTAGAAATTCACCTACTGAGGCGACACTCAAGTTGAATCGATCGAAGAGGGCCC 378  
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGly 120  
Db 379 GGACAGACAGAAATCACTCCACAGGAAGGAGACTGATACATGCACAAATTCCTGTGAGG 438  
Qy 121 ThrHisGlyGlyValThrGlyThrProIleValIleValProAsnThrAspGln 140  
Db 439 ACACATGAAGGGGTGACTACTGGAACGCCAATTCCTGTATTTGTCGCCAAACACAGATCA 498  
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160  
Db 499 ATAGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGCGCACT 558  
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180  
Db 559 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCCTCCGGGCAGAAA 618  
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200  
Db 619 ACCGTGGAAGGGTGGCTGCGAGGGGCCCTCCCAAGAAAATTCCTTAAGCTCAAAATGTGA 678  
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220  
Db 679 TTAGNGATCTTGCTGTTGTTTCCAAAGTGCATCAGGTGTGCTCCCAAGAGAGCGGGTT 738  
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240  
Db 739 GATTATGGGTCTGTAACTTTCGAACAGATAGAGCAACATCGTTAGATGCTCTGATCCA 798  
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgLysSerVal 260  
Db 799 GAGTACGCAGAGAAAGATGATAGCAATCAGACAGTACAGGTTCGAGGGGGATTCGGTTC 858  
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280

Db 859 GGTGGAGTGATCATCATGCGTTCGTAGAAAAGTTCTCCGCGGGCTCGGTTCTCTGTGTTTC 918  
Qy 281 AsplysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300  
Db 919 GACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTCTATTCTCTGCGAGCAACGGGTTTC 978  
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320  
Db 979 GAGATTGGCAGCGGATTTCGCCCGGACCGACTTGACAGAAAGTGAGCAATAATGATGATTT 1038  
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340  
Db 1039 TATATGATTAAGCTCGAAGTGTCAAGACACGAGCTAATCGCTCGGGTGTGTGCGAGGA 1098  
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360  
Db 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTTAAGCCGACACCATCTATC 1158  
Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380  
Db 1159 GGGGTGAAAACAGAAACACCGTGTCAAGGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGG 1218  
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400  
Db 1219 CGCATGACCCATGCGTCCGCCCTCGAGCTGTTCTGTGGTGGAAATCCATGGCCGCTTG 1278  
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420  
Db 1279 GTCTCATGACACGAGCTGATGGCGCACGTGGCTCAGTGCAGATGTTTCGGGCTCAATACT 1338  
Qy 421 AlaLeuGlnGluProValGlySerPhe 429  
Db 1339 GCACTTCAAGAACCGATTGGCTCTTTTC 1365

## RESULT 2

US-10-425-115-138337  
; Sequence 138337, Application US/10425115  
; Publication NO. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 138337  
; LENGTH: 1845  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_57640C.1  
US-10-425-115-138337

Alignment Scores:  
Pred. No.: 1,09e-226 Length: 1845  
Score: 2090.00 Matches: 422  
Percent Similarity: 84.80% Conservative: 2  
Best Local Similarity: 84.40% Mismatches: 5  
Query Match: 95.00% Indels: 71  
DB: 20 Gaps: 2

US-10-677-179-8 (1-429) x US-10-425-115-138337 (1-1845)

Qy 1 MetThrThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20  
Db 111 ATGACGACCGTGGCCACCGCCACAGCAGGTGGCGCACTCACGGGCACGGCTCGCACCCCGC 170  
Qy 21 AlaileGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40

Db 171 GGCATCGGCGCTTGTGAGTTTGGCCGAGCTCTCTCTCCCTCGCTTCGCGTGCAC 230  
Qy 41 ArgCysArgThrAlaArgLeuGluValIleAlaSerGlyAsnThrPheGlyAsnTyrPhe 60  
Db 231 CGCTGCGGCACTGCTCGCTAGAGGTGAAGGATCTCGGAAACACGTTTGGAAACTACTTT 290  
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyGlyValGlyCysValIleSerGlyCys 80  
Db 291 CAGGTTGCAACCTATGGTGAATCTCATGGGGTGGTGTGGTTGTATCATGTGGTGT 350  
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100  
Db 351 CCACCTAGAAATTCCTACTCAGGCGAGACCTTACAAGTTGAATCGATCGAAGACGGGCC 410  
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysValIleLeuSerGly 120  
Db 411 GGTCAAGAGAAATACCTCCCAAGAAAGAGAGATGATACATGCAAAATTCGTGAGGG 470  
Qy 121 ThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGln 140  
Db 471 ACACATGAAGGGGTGACTACTGTAGCCCAATCTTGTATTGTCTCCAAACACAGATCA 530  
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160  
Db 531 ATAGGCACTGATCACCGTGAAATAGCAATGTGTACCGACCTTCTCATGCGACGCAACT 590  
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180  
Db 591 TATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGGAGGTGGAGGTGCTCGGGCAGAGA 650  
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200  
Db 651 ACCGTTGAGGTTGGCTGCGAGGGCCCTCGCAAGAAATTTCTTAAGCTCAATGTGA 710  
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220  
Db 711 TTAGAGATCTTGTCTGTTGTTTCCAAAGTGATCAGGTTGTGCTCCAGAGAGACGGGT 770  
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240  
Db 771 GATTATGGTCTGTAACTTGGAAATATAGAGACCAATCATGTTAGATGCTCTGATCCA 830  
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgLysAspSerVal 260  
Db 831 GAGTACGACAGAGAATGATAGACCAATCGACAGATACGAGTTTCAGAGGGAATTCGGTC 890  
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280  
Db 891 GGTGAGTGTATCATGCTGCTAGAAAGCTTCCTCGCGGGCTCGGTTCTCTGTGTTTC 950  
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300  
Db 951 GACAAAGCTCGAATCGGAATCGCAAAAGCTATGCTTCTATTCTCGCAGCAACGGGTT 1010  
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320  
Db 1011 GADATGGACGGGATTCGCGGGACCGACTTGACAGGAAGTGACATAAGTATGATGATT 1070  
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGly 337  
Db 1071 TATATGGATAGGCTTGAAGTGTACAGACACGAGCAATATCGCTCGGGGGTGTGAGGTT 1130  
Qy 337 337  
Db 1131 TTTTTCCTTCATCTTAATAAAACCCCTATCGATTGTGACTGTGCAATGTTCTCTAAAA 1190  
Qy 338 -----ValGlnGlyGlyIleSerAsnValGluIleVal 348  
Db 1191 AAAATGGAAATCGTTTGTGTTGTAATGTGCGAGGAGGGATATCGAATGGTGAGATTGTG 1250  
Qy 349 HisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSer 368

Db 1251 CACTTCAAAGTTGCTTTTAAAGCGCACCATCTATCGGGGTGAAACAGACACACTGTGTCA 1310  
Qy 369 ArgGluArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaPro 388  
Db 1311 AGGAGAGCGTCAGAACGTTGAGCTTTTGGCAAGAGGGCGCCATGACCCATCGTCCGCCCT 1370  
Qy 389 Arg 389  
Db 1371 CGAGGTAAATGCGACCGTGTAAATGACGCTCCCTCGCTGTAGTGTCTTACCATCTTTT 1430  
Qy 390 -----Al 390  
Db 1431 ATTTCTTCCCGCGACTCGGACTCGATATGAGGTTTGTCTCTTACGCTGGCTTCTGCGAGC 1490  
Qy 390 aValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVa 410  
Db 1491 TGTCTCTGTGTGAAATCCATGGCGCGCTTGGTCTCTGTGACCAAGCTGTATGGCGCAGT 1550  
Qy 410 LalaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProValGlySerPhe 429  
Db 1551 GGCCCAAGTGCAGATGTTGGCGCTCAATGTGCTCACTTCAAGAACCAAGTTGGCTCTTTC 1608

## RESULT 3

US-10-425-114-30604  
; Sequence 30604, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 30604  
; LENGTH: 1496  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73045G02\_FLI  
US-10-425-114-30604

Alignment Scores:  
Pred. No.: 1,82e-218 Length: 1496  
Score: 2016.50 Matches: 398  
Percent Similarity: 93.01% Conservative: 1  
Best Local Similarity: 92.77% Mismatches: 3  
Query Match: 91.66% Indels: 27  
DB: Gaps: 1

US-10-677-179-8 (1-429) x US-10-425-114-30604 (1-1496)

Qy 1 MetThrThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20  
Db 68 ATGACGACCGTGCCTCAAGCCACAGCAGGTGGCGCAGCTCACGGGACGCTCGCACCCGCG 127  
Qy 21 AlaIleGlyAlaLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHis 40  
Db 128 GCGATCGCGCTTGTGTGGAGTTTGGCCCGACGCTCTCTCCCTCGCGTTCGCGGTGAC 187  
Qy 41 ArgCysArgThrAlaArgLeuGluValIleAlaSerGlyAsnThrPheGlyAsnTyrPhe 60  
Db 188 CGCTGCGCACTGCTCGCTAGAGTGAAGGCACTCGAAGACACGTTTGGAACTACTTT 247  
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCys 80  
Db 248 CAGGTTCGAACCTATTGTGAATCTCATGGGGGTGGTGTGGTTGTGTATCAGTGGTGT 307

```
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
Db |||||
Qy 308 CCACCTAGAAATCCACTCACTGAGGCAGACCTACAGTTGAACCTCGATCGAAGACGGCCC 367
Db |||||
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysValLeuSerGly 120
Db |||||
Qy 368 GGACAGACAGAGAAATAACCTCCACAGGAAGAGAGCTGATACATGCMAAATTTCTGTGAGG 427
Db |||||
Qy 121 ThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
Db |||||
Qy 428 ACACATGAAGGGGTGACTACTGGAAGCGCAATCTTGTATTGTGCCAAACACAGATCAA 487
Db |||||
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160
Db |||||
Qy 488 ATAGGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGCGCACT 547
Db |||||
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
Db |||||
Qy 548 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTGTGCGGACAGAA 607
Db |||||
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysCysGly 200
Db |||||
Qy 608 ACCGTGGAAGGGTGGCTGCAGGGGCCCTCGCCAGAAAAATTTCTTAAGCTCAAAATGTGGA 667
Db |||||
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
Db |||||
Qy 668 TTAGAG----- 673
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
Db |||||
Qy 674 -----ATAGAGACAACATCGTTAGATGCTCTGATCCA 706
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgLysSerVal 260
Db |||||
Qy 707 GAGTACGAGAGAAAGATGATGACCAATGCACAGAGTACGAGTTCGAGGGGATTCGGTTC 766
Db |||||
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
Db |||||
Qy 767 GGTGGAGTGATCACATGGCTGCTAGAAAAGTTCTCTCGGGGCTCGGTTCCTGTGTTC 826
Db |||||
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db |||||
Qy 827 GACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCTCGCGAGCAACGGGTTTC 886
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
Db |||||
Qy 887 GAGATTGGCAGCGGATTCGCGGGGACCGACTTGACAGGAAGTGAGCATAAATGATGAGTTT 946
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
Db |||||
Qy 947 TATATGGATNAGGCTGGAGGTGTACAGACAGGACTAATCGCTGGGTGGTGTGCGAGGGA 1006
Db |||||
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
Db |||||
Qy 1007 GGGATATCGAATGGTGAGATTGTGCATTTCAAAAGTTGCTTTTAAGCCGACCACTCATC 1066
Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
Db |||||
Qy 1067 GGGGTGAACAGAACACCGTGTCAAGGGAGCGGTCAAGACGTTGAGCTTCTAGCAAGAGGG 1126
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
Db |||||
Qy 1127 CGCCATGACCCATGGTGGCCCTCGAGCTGTTCTGTGTGGTGAATCCATGGCGCGTTG 1186
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
Db |||||
Qy 1187 GTCTCTCATGGACAGCTGATGGCGCACGTGGCTCAGTGCAGATGTTTCGCGCTCAATACT 1246
Qy 421 AlaLeuGlnProValClySerPhe 429
Db |||||
Qy 1247 GCACCTTCAAGAACCCAGTTGGCTCTTTTC 1273
```

```
US-10-425-115-138335
; Sequence 138335, Application US10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 138335
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57639C.1
US-10-425-115-138335

Alignment Scores:
Pred. No.: 7,03e-218 Length: 1541
Score: 2011.50 Matches: 397
Percent Similarity: 92.77% Conservative: 1
Best Local Similarity: 92.54% Mismatches: 4
Query Match: 91.43% Indels: 27
DB: Gaps: 1

US-10-677-179-8 (1-429) x US-10-425-115-138335 (1-1541)
Qy 1 MetThrThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
Db |||||
Qy 68 ATGACGACCGTGGCCACAGCCACAGAGGTGGCGCATCTCACGGGCACGGCTCGCACCCCGC 127
Db |||||
Qy 21 AlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHis 40
Db |||||
Qy 128 GCGATCGGCGCTTGTGTGGAGTTTGGCCAGCCTCTCTCTCCCTCCGCTTCGCGGTGCAC 187
Db |||||
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60
Db |||||
Qy 188 CGCTGCGGCATCTGCTCCCTAGAGGTGAGGCATCTGGAACACAGTTTGGAAACTACTTT 247
Db |||||
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyGlyValLysValLysSerGlyCys 80
Db |||||
Qy 248 CAGGTTGCAACCTATTGTGAATCTCATGGGGGTGGTGTGGTTGTGTATCAGTGGTTGT 307
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
Db |||||
Qy 308 CCACCTAGAAATCCACTCACTGAGGCAGACCTACAGTTGAACCTCGATCGAAGACGGCCC 367
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGly 120
Db |||||
Qy 368 GGACAGACAGAGAAATAACCTCCACAGGAAGAGAGCTGATACATGCMAAATTTCTGTGAGG 427
Qy 121 ThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
Db |||||
Qy 428 ACACATGAAGGGGTGACTACTGGAAGCGCAATCTTGTATTGTGCCAAACACAGATCAA 487
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160
Db |||||
Qy 488 ATAGGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGCGCACT 547
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
Db |||||
Qy 548 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTGTGCGGACAGAA 607
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysCysGly 200
Db |||||
Qy 608 ACCGTGGAAGGGTGGCTGCAGGGGCCCTCGCCAGAAAAATTTCTTAAGCTCAAAATGTGGA 667
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
```

```
Db 668 TTAGAG----- 673
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
Db 674 -----ATAGAGCAACATCGTTAGATGTCCTGATCCA 706
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260
Db 707 GAGTACGCAGAGAGATGATAGCGCAATCGACAGAGTACGAGTTCGAGGGATTCGGTC 766
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
Db 767 GGTGGAGTATCAGATCGCTCGCTAGAAACGTTCTCTCGCGGGCTCGGTTCTCTGTGTTC 826
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db 827 GACAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTCGAGCAACCGGGTTC 886
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluLysAsnAspGluPhe 320
Db 887 GAGATTGGCAGCGGATTCCCGGGACCGACTTGACAGGAGTGCAGCAATATGATGATTT 946
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
Db 947 TATATGGATAGGCTGGAAAGTGTTCAGGACACGAGCTAATCGCTCGGGTGGTGCAGGGA 1006
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
Db 1007 GGGATATCGAATGGTGAGATTGTGCATCTCAAGTGTCTTTTAAAGCCGACACCATCTATC 1066
Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
Db 1067 GGGGTGAAACAGACAAACCGTGTCAAGGGAGCGTCAAGAGCTTGAGCTTCTGCAAGAGGG 1126
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
Db 1127 CGCCATGACCCATGCGTCCGCCCTCGAGCTGTCTCGTGTGGTGGATCCATGCGCGGTG 1186
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
Db 1187 GTGCTCATGGACCGCTGATGCGCGCAGCGTGGCCAGTGCAGATGTTGCGCGTCAATGCT 1246
Qy 421 AlaLeuGlnGluProValGlySerPhe 429
Db 1247 GCACCTTCAGAACCAAGTTGGCTCTTTTC 1273
```

## RESULT 5

```
US-10-425-115-171490
; Sequence 171490, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 171490
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRP4577_87981C.1
US-10-425-115-171490
```

## Alignment Scores:

```
Pred. No.: 8,17e-194 Length: 1792
Score: 1800.50 Matches:
```

```
Percent Similarity: 89.44% Conservative: 31
Best Local Similarity: 82.16% Mismatches: 42
Query Match: 81.84% Indels: 3
DB: 20 Gaps: 2
```

US-10-677-179-8 (1-429) x US-10-425-115-171490 (1-1792)

```
Qy 4 ValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArgAlaIleGly 23
Db 178 GTGTCCGAGCGCGGTGTCCGCCAGGGCGTCCACACGGTTTCTCCCGCGGGATAGGC 237
Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaVal---HisArgCys 42
Db 238 GCGCTCCCGAGTCCGCC---ACGTCCTCCGCTTATCGTCGCGCGCGTCCGC 291
Qy 43 ArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnVal 62
Db 292 CGCGCGCGCGCTAGAGGTGAAGGCATCGGGAAATGTGTTCCGGGAACCTACTTCCAGGTT 351
Qy 63 AlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProPro 82
Db 352 GCACCTATGGCAATCCCATGGAGGGGGTGTGGTTGCGTTATCGATGGCTGCCACCC 411
Qy 83 ArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGln 102
Db 412 AGAATTCCTCTCACTAGGCGAGACATCGAAGTAGAAGTCCGATAGAAGACGTCCTCGGTC 471
Qy 103 SerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHis 122
Db 472 AGTAGAATTAACACCCCAAGAAAGAGAGACTGATCATGCAAAATTTCTATCAGGGACAT 531
Qy 123 GluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGly 142
Db 532 GATGGAGTACTCTGTGTACCAATTCACGCTTTGTCCCAACACAGATCAAGGGGT 591
Qy 143 SerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAsp 162
Db 592 GGTGATTACAGTGAATGCTGAAGCGGTACAGACCATCCCATGCAGATGCAACCTATGAC 651
Qy 163 PheLysTyrGlyValArgAlaValGlnGlyArgSerSerGlyArgLysThrVal 182
Db 652 TTCAAGTATGAGTTAGAGCTGTGTCAGGGAGGTGGAAGTGCATCAGCCAGAGAAACCAT 711
Qy 183 GlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGlu 202
Db 712 GGCAGGGTGGCTCGAGGAGCTCTTGCAAGAAATTTCTAAAGCTCAAAATCAGGAGTGAG 771
Qy 203 IleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyr 222
Db 772 ATCTTGGCAATTTGTTCTAAAGTGCACCAAGTCGTAATCCAGAAAGATGCAGTTGATAT 831
Qy 223 GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr 242
Db 832 GAGACTGTAACTTTGGAACATATATAGAGAGCAACATCGTTAGATGTCCTGTATCCAGAA 891
Qy 243 AlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGly 262
Db 892 GCAGAGAGATGATGCTGCCATTTGATACGGTACAGTTAGGAGATTCATTTGGTGGG 951
Qy 263 ValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLys 282
Db 952 GTGCTCATCATGCAATTCGAAGAAATGTTCTCTGCTGTGCTCTCTGCTCTCTGCTTTGACA 1011
Qy 283 LeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIle 302
Db 1012 CTTGAAGCTGAACCTGGCAAGGCCATGCTTTCTCTCTCTGCAAGCAAGGGGTTTGAGAT 1071
Qy 303 GlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMet 322
Db 1072 GGCAGTGGGTTGCTGTGTGACGACTTTTACTGGAAGTGAGCATATATGATGAGTTCTATATG 1131
Qy 323 AspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyIle 342
```



Db 1132 GATGAGGCTGGAATGTGAGGACAGCAACTAATCGCTCAGCGGTGTTTCAGGAGGGGATA 1191  
 Qy 343 SerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362  
 Db 1192 TCAATGTGTGAATTAATTTACTTCAAGTGGCTTTTAAGCCACAGCAACTATCGGAAG 1251  
 Qy 363 LysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaAatGlyArgHis 382  
 Db 1252 AAGCAAAATACGTGTCTCAAGGAGCATGAGGATGTGAACATTTTGGCAAGGGGGCGCCAT 1311  
 Qy 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402  
 Db 1312 GACCCCTGTGTGTCCTCGAGCTGTTCTATGTGTGAATCCATGGCTGCGTGTCTG 1371  
 Qy 403 MetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeu 422  
 Db 1372 ATGGACCAAGCTCATGGCGCATATTGCCCGCATATTGCCGATGTGAGATGTTTCGGCTGAACCTTGGCCCTA 1431  
 Qy 423 GlnGluProValGlySer 428  
 Db 1432 CAAGAGCCCATTTGGCTCT 1449

## RESULT 6

US-10-677-179-1  
 ; Sequence 1, Application US/10677179  
 ; Publication No. US20040082050A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cahoon, Rebecca E.  
 ; APPLICANT: Falco, Saverio Carl  
 ; APPLICANT: Pember, Stephen O.  
 ; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes  
 ; FILE REFERENCE: BB-1159-A  
 ; CURRENT APPLICATION NUMBER: US/10/677,179  
 ; CURRENT FILING DATE: 2003-10-02  
 ; PRIOR APPLICATION NUMBER: US/09/743,207  
 ; PRIOR FILING DATE: 2001-01-04  
 ; PRIOR APPLICATION NUMBER: PCT/US99/16353  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/093,611  
 ; PRIOR FILING DATE: 1998-07-21  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 1  
 ; LENGTH: 1635  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 US-10-677-179-1

Alignment Scores:  
 Pred. No.: 1,578-193 Length: 1635  
 Score: 1797.50 Matches: 349  
 Percent Similarity: 89.44% Conservative: 32  
 Best Local Similarity: 81.92% Mismatches: 42  
 Query Match: 81.70% Indels: 3  
 Gaps: 2

US-10-677-179-8 (1-428) x US-10-677-179-1 (1-1635)

Qy 4 ValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArgAlaIleGly 23  
 Db 64 GTGTGCGAGCGCGGTGTCCGCCAGGGCGTCCACACGGTTTCTCCCGCGGGATAGGC 123  
 Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaVal ---HisArgCys 42  
 Db 124 GCGCTCCGGAGTCCGCCGCC ---ACGTCCCTCCGGTTATCCGTGCGCGCGCGTGC 177  
 Qy 43 ArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPheGlnVal 62  
 Db 178 CCGCGCTCACGCTAGAGGTGAGGCATCAGGAATGTGTCGGGAATCTTTCAGGT 237  
 Qy 63 AlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProPro 82  
 Db 238 GCAACCTATGGCAATCCATGGAGGGGTGTGTGCTTCGGTTATCAGTGGCTGCCACCC 297

Qy 83 ArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGln 102  
 Db 298 AGAATTCCTCTCACTGAGGAGACATCGAAGTAGAACTCGATAGAAAGACGTCGGGGTCAA 357  
 Qy 103 SerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHis 122  
 Db 358 AGTAGAATTAACAACCCNAGAAAGGAGACTGATACATGCAAAATTTCTATCAGGGACACAT 417  
 Qy 123 GluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGly 142  
 Db 418 GATGGATGACTACTGTGTACCAATTCACGTCTTTGTGTCACCAACACAGATCAAGGGGT 477  
 Qy 143 SerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAsp 162  
 Db 478 GGTGATTTACAGTAGTAATGTCTAAGGGGTACAGACCATCCCATCAGATGCAACCTATGAC 537  
 Qy 163 PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal 182  
 Db 538 TTCAAGTATGGAGTTAGAGCTGTGCAGGAGGTGGAAAGTTCATCAGCCAGAAACCATTT 597  
 Qy 183 GlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLysCysGlyLeuGlu 202  
 Db 598 GGCAGGGTGGCTGCAGGAGCTTTGCAAAAGAAAATCTAAAGCTCAAAATCAGGAGTGGAG 657  
 Qy 203 IleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyr 222  
 Db 658 ATCTTGGCATTTGTTTCTAAAGTGCACCAAGTCGTACTTCCAGAAGATGTCAGTTGATAT 717  
 Qy 223 GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr 242  
 Db 718 GAGACTGTAACTTGGNACATATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAATAT 777  
 Qy 243 AlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGly 262  
 Db 778 GCAGAGAAGATGATTTGCTGCCATTCATCGGTACGAGTTCAGAGAGATTCAAATTTGGTGG 837  
 Qy 263 ValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLys 282  
 Db 838 GTCGTCACATGCATTCGAAAGATTTCTCTCGTGGTCTTGGCTCTCTCTGTTTTTGACAA 897  
 Qy 283 LeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIle 302  
 Db 898 CTTGAAGCTGAACCTGGCAAAAGCCATGCTTCTTCTTCTCCAGCAAGAGGGGTTTGAGATT 957  
 Qy 303 GlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMet 322  
 Db 958 GGCAGTGGGTTCGCTGCTACCGACTTTTACTCGAAGTGAGCATTAATGATGATTTCTATATG 1017  
 Qy 323 AspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIle 342  
 Db 1018 GATGAGGCTGGAATGTGAGGACACGAACATAATCGCTCAGGCGGTGTTCAGGAGGGGATA 1077  
 Qy 343 SerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362  
 Db 1078 TCAATGGTGAATTAATTTACTTCAAGTGGCTTTTAAGCCACACAGCAACTATCGGAAG 1137  
 Qy 363 LysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArgHis 382  
 Db 1138 AAGCAAAATACTGTGTCAAGGAGCATGAGGATGTTGAACTTTTGGCAAGGGGGCGCCAT 1197  
 Qy 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402  
 Db 1198 GACCCCTGTGTGTCCTCAGAGCTGTTCTATGTGTGAATCCATGCGTGGCTGGTCTGCTG 1257  
 Qy 403 MetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeu 422  
 Db 1258 ATGGACCAAGCTCATGGCGCATATTGCCAGTGTGAGATGTTTCGCTGAACCTTGGCCCTA 1317  
 Qy 423 GlnGluProValGlySer 428  
 Db 1318 CAAGAGCCCATTTGGCTCT 1335

US-10-425-114-32244

**Alignment Scores:**

US-10-677-179-8 (1-429) X US-10-425-114-32244 (1-1603)

## RESULT 8

```

US/10437963
; Sequence 99207, Application US/10437963
; Publication No. US2004012334A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalick, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Zuo, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid
; TITLE OF INVENTION: Plants and Uses T
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 99207
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Oryza sativa

```





820 CATAAAGATGTCTTCAGAGGACCTTATTGATCATGACACTCTGACTCTTGATCAGATT 879

231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250

880 GAGAGTAACATTGTTTCGATGTCACAGACCCGGAGTATGACAGAGAGATGATATCTGCATTT 939

251 AspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsn 270

940 GATGCTGTGCGAGTGAGAGGTGATTCCTGTTGGTGGTGTGTGACATGCAATTTGTGAGGAGAAC 999

271 ValProArgGlyIleGlySerProValPheAspLysIleGluSerGluLeuAlaLysAla 290

1000 TGTCCACGAGGTCTCGGTTCACCAAGTATTTGACAACTTGAAGCTGAGCTGGCTAAAGCT 1059

291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310

1060 GCAATGTCATGTCCTGCACCAAGGCTTTCAGTTGGTAGTGGGTTTGAGGACCCCTTT 1119

311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330

1120 TTGACTGGAGTGAACAATGATGAGTTCCTATATAGATGAACATGGAACACACAGAACA 1179

331 ArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPhe 350

1180 AGAACAAATCGCTCTGGTGGGATACAGGGTGGAAATTTCCAATGGGGAATCATTAATATG 1239

351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370

1240 AGAATAGCTTTCAGCCCAACATCAACAAATTTGGAAGAAGCAAAAGACTGTGACTCGAGAT 1299

371 ArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390

1300 AAAAAAGAAACAGAGTTTATAGCCGCTGGTCCCATGATCTCTGTGTGTGCCAAGAGCT 1359

391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410

1360 GTACTCTATGTTAGAACCAATGGTAGCTTTGGTCTTCTGTGGACCAATTTGATGGCACAATAT 1419

411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426

1420 GCGCAGTGTAACTCTTTTCCGTAACACTCAGATTTCAGAAACCCCTTG 1467

RESULT 11

US-10-696-616-27

; Sequence 27, Application US/10696616

; Publication No: US20040191850A1

; GENERAL INFORMATION:

; APPLICANT: Levin, Joshua

; APPLICANT: Tossberg, John

; APPLICANT: Zhou, Qing

; APPLICANT: McElver, John

; APPLICANT: Aux, George

; APPLICANT: Budziszewski, Greg

; APPLICANT: Thomas, Carla

; APPLICANT: Patton, David

; APPLICANT: Frye, Catherine

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Growth

; TITLE OF INVENTION: Development and Uses Thereof

; FILE REFERENCE: 70160USNP

; CURRENT APPLICATION NUMBER: US/10/696.616

; CURRENT FILING DATE: 2003-10-29

; PRIOR APPLICATION NUMBER: US 60/423,519

; PRIOR FILING DATE: 2002-11-04

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 27

; LENGTH: 1311

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1311)

; OTHER INFORMATION:

309	ATGCAAGTAGAACTCGACCCGAGAGCGCCAGCGCCAGACGAGAAATAAACCAACCCCAAGAAAG	368
111	GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro	130
369	GAGACTGACACTTGC AAAATCTTTACAGGCACACATGAAGGAAATGACCACTGGGACACCA	428
131	IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150
429	ATTCAATGTTTTTGTCGCCGAACACACAGATCAGACAGGGGTGATTACAGTGAATGGCTAAG	488
151	ValTyrrArgProSerHisAlaaspAlaThrTyrrAspPheLysTyrrGlyValArgAlaVal	170
489	GCCTACAGACCTTCACATGCAGATGCAACTTATGACTTCAAATACGGTGTTAGAGCAGTG	548
171	GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu	190
549	CAGGGAGGTGGAAGATCATCAGCAAGACAGACCACTTGGAAAGGTGGCTGCAGGAGCTCTT	608
191	ProLysIlyIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal	210
609	GCAAGAAGAAAATTTCTTAAGCTCAAATCTGGAGTAGAGATCTTTGGCGCTTTGTGTCGAAGGTG	668
211	HisGlnValValLeuProGluAspAlaValAspTyrrGlySerValThrLeuGluGlnIle	230
669	CATCAAGTTGTACTACCCAGACATGCCCTGTGATTATGACACTGTAAACATGGAACAGATA	728
231	GluSerAsnIleValArgCysProAspProGluTyrrAlaGluLysMetIleAspAlaIle	250
729	GAAAGCAACATTGTTAGATGCTCGATCCAGAATATGCACAGACATGATTGATGCACTC	788
251	AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn	270
789	GATAAAGTACAGATTAGAGGTGATTCGATTGGTGTGGTGCACATGCATTGCCAAGAAT	848
271	ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla	290
849	GTTCCTCGTGGATTGGCTTCCTCTGATTTCACAAACTTGAGGCTGAATTGCCGAAGCT	908
291	MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr	309
909	ATGCTTTCTCTTCGCAAGCAAGGGGTTTCAGATCGCAGTGAATTTGTGTTCACT	965

## 5

609	Db	GCAAGAAAAATCTTAAAGCTCAAAATCTGGAGTAGAGATCTTTGGCGCTTTGTGTCTCAAGGTTG	666
211	Qy	HisGlnValValLeuProGluuAspAlaValAspTyrGlySerValThrLeuGluGlnIle	230
669	Db	CATCAAGTTGTACTACACAAGATGCCCTTGATTATGACACTGTAACTATGAAACAGATA	728
231	Qy	GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle	250
729	Db	GAAGAGCAACATGTTAGATGTCCTGATCCAGATATGCACAGAAAGATGATTGATGCACATC	788
251	Qy	AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaAlaArgAsn	270
789	Db	GATAAGTACGAGTTAGAGGTGATTCGATTGCTGGTGGTCACATGCATTCGACAAGAAT	848
271	Qy	ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla	290
849	Db	GTTCCTCGTGGGATTGGCTCTCCTGTATTTGACAAACTTGAGGCTGAATTGGCGAAAGCT	908
291	Qy	MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr	309
909	Db	ATGCTTTCTCTCTCGAAGCAAGGGGTTTGAGATCGGCAGTGGGATTTGTGTTCTACT	965

RESULT 13  
US-10-369-493-26364  
U

```

: Sequence 26364, Application US/10369499
: Publication No. US20030233675A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
:

```

;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

```

, FILE REFERENCE: 38-10(52052)B
, CURRENT APPLICATION NUMBER: US/10/369,493
, CURRENT FILING DATE: 2003-02-28
, PRIOR APPLICATION NUMBER: US 60/360,039
, PRIOR FILING DATE: 2002-02-21
, NUMBER OF SEQ ID NOS: 47374
, SEQ ID NO 26364
, LENGTH: 1089
,

```

```

?      TYPE: DNA
?      ORGANISM: Synechocystis sp.
US-10-369-493-26364

Alignment Scores:
Pred. No.:      6 32e-125      Length:      1089
Score:          1131.50      Matches:      225
Percent Similarity: 77.90%      Conservative: 57
Best Local Similarity: 72.15%      Mismatches:  77
Query Match:     54.16%      Indels:      3
DB:              17      Gaps:        1

```



US-10-677-179-8 (1-429) x US-10-369-493-42555 (1-1089)

Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72  
Db 4 GAAACACGTTGGCAGTTTATTTGCGATTAACCACTTCGGTGAGTCCCATGGGGGGGC 63

Qy 73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92  
Db 64 GTGGGGGTAATATTGATGGTGGCTCCCGATTTGGAATTTCCCCAGAGGAATCCAA 123

Qy 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112  
Db 124 GTAGACCTAGACCGCGCGCCAGCAGACAGCAAAATTAACACCCCTCGTAAGAGCA 183

Qy 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132  
Db 184 GATCAATGCAAAATTTCTCTGGCGTGTGAGGGTAAACTCTGGGTACACCGATGCC 243

Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152  
Db 244 ATTTTGGTGGAAATAAGATACCTCGTCCCAAGATTATATGATGAGATGGCGCAAAATGAT 303

Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172  
Db 304 CGTCTTCCACGCGATGCCACCTATGACGGAATATGCGATTCGCAATTGGCAGGA 363

Qy 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192  
Db 364 GGAGGTGCGCTCTTCTGTAGGAAACCACTTGGCGGGTGGCGCCAGGGGGCGATGCCAAA 423

Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212  
Db 424 AAAATTCCTCCGCAAGTTGCTAAATGATAGGCTATAGGCTACGTCACCAACGTCACAAAGAC 483

Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232

US-10-677-179-8 (1-429) x US-10-369-493-42555 (1-1086)

Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72  
Db 4 GGCAATACTTTGGACATCTGTTTCGGATCACTACTTTTCGGAGTCAACGGTGGAGGC 63

Qy 73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92  
Db 64 GTGGGGTGTGATTTGCTGGGGTATATGAAGGCAAACTCTCGTGAAGAAATTCAA 123

Qy 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112  
Db 124 CTGAGTTAGACCGCAGACGACCCGACAGAGTAGATTACAGCCCTCGTAAGAGCA 183

Qy 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132  
Db 184 GACACCTGTGAGATTTTGTCTGGGGTATATGAAGGCAAACTCTCGTGAAGAAATTCAA 243

Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152  
Db 244 ATTTTGTGCGAAATAAGATACCTCGTCCCAAGATTATATGATGAGATGGCGCAAAATGAT 303

Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172  
Db 304 CGTCTTCCACGCGATGCCACCTATGACGGAATATGCGATTCGCAATTGGCAGGA 363

Qy 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192  
Db 364 GGCGGTAGATCTCTCAGCGCGGAGACAATTTGGTAGAGTAGCCGCGGTGCGATCGCTAAA 423

Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212  
Db 424 AAAATTCCTCCGCAAGTTGCTAAATGATAGGCTATAGGCTACGTCACCAACGTCACAAAGAC 483

Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232



```

Db 484 TTG-----GAAGGTGTTGTAGATCCCAATACCGTCACCTAGACCAAGTGGAAAGC 534
Qy 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db 535 AATATCGTTGCGTTCCTGATCGCGAATGGCAGACCGCATGATGAATTAATCGAACAA 594
Qy 253 ValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValPro 272
Db 595 ACAGGTAGACAAAGTGATCTATCGCGGTGTAGTGGAAATGGTGGGCGGAAATGTTCT 654
Qy 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db 655 AAAGGCTTGGGTGAACCTGTATTTCGATAAGTTGGAAGCAGATATCGCAAGCTGTGATG 714
Qy 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 715 TCTTACCTGCTGAGCAAGGTTTGAATTCGGTTTCAGGTTTTCGCGGACTCTGCTAAC 774
Qy 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 775 GGATTTGAGCATAAACGATATATATATGATGAACCGTGAGATTGCGACAGTAAT 834
Qy 333 AsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValIlePheLysVal 352
Db 835 AACCGTTCGGTGGGATTCAAGCGGTATTGCAACGGAGAGAAATCATTTTTCGAGTT 894
Qy 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
Db 895 GCATTTAAACCCACACGACCGATAGAAAGAGCAAAACACAGTTACTCTGTAAGCGCAA 954
Qy 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaValPro 392
Db 955 GAAACCTATTAGCCGCCAAGGTAGACATGATCCCTGTGTTCTACCCCGTCAGTCCCG 1014
Qy 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db 1015 ATGTTAGAAAGCAATGGTGGCGTGTGTTACTGTGCGATCATTTGTTACGCCATCATGGACAG 1074
Qy 413 CysGluMet 415
Db 1075 TGTAAAGTG 1083

```

## RESULT 15

```

US-10-369-493-43897
; Sequence 43897, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cad, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43897
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-43897

```

## Alignment Scores:

Pred. No.:	6.3e-124	Length:	1086
Score:	1182.50	Matches:	220
Percent Similarity:	78.51%	Conservative:	65
Best Local Similarity:	60.61%	Mismatches:	75
Query Match:	53.75%	Indels:	3

```

DB: 17 Gaps: 1
US-10-677-179-8 (1-429) x US-10-369-493-43897 (1-1086)
Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72
Db 4 GGCATAATATTTTGGTTCATTTATTTTCGATTAGTACTTTTGGCGAGTCTTCACGGCGCGGT 63
Qy 73 ValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGln 92
Db 64 GTGGGGTGTGATGTGTTGCTCCCAACATAGAAATTTCCGCAAGAAATTTAG 123
Qy 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
Db 124 GTAGAACTAGATAGAGCGCGCGGCAAAAGTAAATTTACGACTCCCGCCAGAAAGCT 183
Qy 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeu 132
Db 184 GATACCTCGAGATTATCTCAGAGATATTTGAAGCGCAAAACGCTGGGAACCCCTATAACG 243
Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
Db 244 ATTTTGGTACGTAATCAAGATACCTCGTCCCAAGATTATGACGAGATGGCACAGAAAGTAT 303
Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db 304 CGGCTTCTCAGCGGATGCACCTATGATGCAAAATATGCGCATTCGCAATTGGCAAGGT 363
Qy 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
Db 364 GGGGTAGGTCTGTCAGCACGTCGAGACAATCGGAGAGTAGCAGCAGGTGCGATCGCTAAA 423
Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
Db 424 AAAATTCCTCGCAAGTTGCCAATGTCGAATATTGCTTACGTTAAGCGGATTAAGAC 483
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db 484 TTG-----GAAGGTGTTAGTGCATCTTAATATCTGTAACTTGAACACAGTAGAAGC 534
Qy 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db 535 AATATCGTTCGTCGCGATGCGAATGTGCGCATGCGCATGATGAATGATTGATGAGCAA 594
Qy 253 ValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValPro 272
Db 595 ATAGGTAGACAAAGCGATTCTATCGCGGTGTCTAGAAATGTGTGGCGCGAAATGTATCCG 654
Qy 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db 655 AAAGTTTTGGCGCAACCGATATTTGATTAATTAGAAGCTGATATCGCTAGGGTGTCTATG 714
Qy 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 715 TCTCTGCTGCTAGCAAAAGCGCTTTGAAATGTTTCCGTTTTCGCGGAACGCTACTAACG 774
Qy 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 775 GGAATTGAGCATACGACGAATTTTATATCATGATCAAAATGATGAATTTCCGCTACAGTAA 834
Qy 333 AsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
Db 835 AACCGTTCGCGGTATTTCAAGCGCGGATTTCTTACCGCGCAAAATATCATTTTTCGAGTT 894
Qy 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
Db 895 GCATTTAAGCGCAAGCAACAAATTTAGAAAGAAAGACAGAAAGCTGTAACTCGTGGCGCAA 954
Qy 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaValPro 392
Db 955 GAAACCTATTAGCAGCAAAAGGACGACGATCTTTGTTGTTATTCGCGGTGCGAGTGCCA 1014
Qy 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412

```

Db           :::|||||:::||||||| ||| ||| ||| |||  
1015 ATGTTTGAGGCAATGGTGGCGTGTACTATGTACCATTTGTTACGCCATCATGGTCAG 1074

QY         413 CysGluMet 415  
            | | : : : : :  
Db         1075 TGCAAAGTC 1083

Search-completed: August 27, 2005, 07:26:24  
Job time : 755 secs

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**